OPTICAL IMAGE ASSESSMENT: a comparative study

by

C.J. Woodruff, B.Sc.(Hons.)

submitted in fulfilment of the requirements

for the degree of

Doctor of Philosophy

University of Tasmania

Hobart

February, 1975

CONTENTS

Page

ABSTRACT

The aberration theory of Buchdahl is extended to allow greater utilisation of it in optical image assessment and automatic design programs. Expressions are presented for the calculation of the derivatives of the seventh order aberration coefficients as well as for the determination of effects on the derivatives of $\bar{O}T$ coefficients due to pupil shifts arising from parameter changes. Where details of the theory for various constructional parameters differ attention has been confined to axial curvature derivatives. The wavefront retardation expansion has been checked for convergence and the results show general agreement with the convergence properties found by other authors for transverse aberration expansions. A series of transformations, valid over the region of convergence of the retardation expansion, is introduced to reduce the exit pupil periphery defined on a reference sphere to a circle. It is shown that, under these transformations, the form of the retardation expansion remains constant and only the coefficients need be altered. These changes are independent of the field angle but depend on the f-number of the system.

A new set of assessment functions, derived from the real and imaginary parts of the optical transfer function, is introduced. It is shown that, in the geometrical limit, these approximate a set of functions defined in terms of the spot diagram distribution. Theoretical and numerical comparisons of these and some other assessment functions are presented. These show that, in general, there is agreement on the ordering of correction states when different criteria are used. However some differences in ranking do arise and these are discussed. It is found that, with a modification to allow for products of two negative quantities, a function based on the variance, $V(r)$, of the aberration difference function provides an extremely versatile assessment quantity. The

usefulness of this new quantity is shown, both theoretically and numerically, to extend far beyond the conditions under which $V(r)$ provides a valid approximation to the modulation transfer function. Predictions of the location of optimum image planes using the various criteria are examined, and it is shown that by choosing different fractional spatial frequencies, r, at which to evaluate $V(r)$, most of these predictions can be obtained using minimization of the variance $V(r)$.

Finally all the assessment functions introduced are used to examine two different processes of automatic optical design based on reduction of transverse aberrations. The first of these is a primitive design program, developed by the writer, which uses derivatives of the spot diagram assessment quantities introduced earlier. The second method, due to Cruickshank, reduces selected aberration coefficients to prescribed residuals determined by the designer on the basis of past experience and the requirements of the user. The results from this work indicate that the choice of assessment function is not critical in attaining an optimum region of parameter space, but the actual optimum point is dependent on the choice of assessment function. The usefulness of the modified function based on $V(r)$ is again illustrated numerically.

Except as stated therein this thesis contains no material which has been accepted for the award of any other degree or diploma in any university, and, to the best of my knowledge and belief, this thesis contains no copy or paraphrase of material previously published or written by another person, except when due reference is made in the text of the thesis.

INTRODUCTION

Optical design has undergone some radical developments over the past twenty years due to two main factors - the introduction of rapid computational facilities, and the development of spatial frequency concepts for image assessment. The first of these developments provided obvious advantages to the optical designer, allowing him to perform rapidly computations which previously would not have been attempted, or at least only as a final stage of image assessment. Such computations included skew ray traces and the calculation of higher order coefficients, as well as derivatives of higher order coefficients with respect to various constructional parameters. However the second of these developments did not provide such obvious advantages, and was not as readily incorporated by designers into their tools of trade. Originally the value of the optical transfer function (o.t.f.) was seen to lie in its representation of the properties of images in which diffraction was significant, this representation being such that information concerning defects in the optical system could be more easily obtained than was the case when only the form of the diffraction image of a point object was available. However spatial frequency functions are not restricted in usefulness to those cases when diffraction is significant and realization of this was shown by the introduction and use of the geometrical transfer function (g.t.f.). This may be calculated from the spot diagram distribution using a fourier transform, whereas the o.t.f. may be calculated from the point spread function by similarly transforming. Other methods of calculation are used, but from the preceding method one can see the relation between the o.t.f. and the spot diagram - both give information regarding the image - forming properties of an optical system for a particular object point. Information regarding the imaging properties of the system for neighbouring points is obtained either by assuming that the aberrations vary negligibly in the neighbourhood of the initial point (the assumption of spatial invariance) or else by calculating further o.t.f.'s or spot diagrams. This assumption of local spatial invariance was, until very recently, implicit in almost all design assessment. Recent work on coherent processing, however, has shown that spatial variance may be utilised under certain conditions, and hence systems designed for such use must be examined with explicit knowledge of their degree of spatial variance.

Most designers accepted the spot diagram as a very useful device for overall image assessment, though the methods to be used for quantitative evaluation of it were not clearly defined. Some designers and theroeticians found the wavefront aberration a very useful quantity, especially if diffraction effects were likely to be significant. In particular the variance of the wave aberration was shown by Marechal to be directly related to the Strehl definition of a system, provided the aberrations were very small. Rayleigh had earlier established his $\lambda/4$ tolerance criterion on wavefront aberration for well-corrected systems. Direct investigation of the effects of aberrations on diffraction images had not resulted in any significant advances in design, so the potential of the o.t.f. appeared to lie in this direction. At this stage a certain amount of work had been done on comparing the effects of using different criteria of image quality on the design. These studies were quite limited and did not consider actual systems, but rather, concentrated on simple combinations of aberration terms. The question as to whether a difference in the design of a system would arise from using different criteria during the design process was not tackled.

The o.t.f. appeared to offer great possibilities as an assessment function in the late 1950's and a very extensive literature developed. However very little appears to have been done during this time to determine whether use of the o.t.f. as an assessment function would affect the design process. A number of approximations to the o.t.f. were introduced by H.H. Hopkins which increased its potential by greatly reducing the computation time involved in getting reasonably accurate o.t.f. values, or at least obtaining spatial frequency information regarding the system.

This thesis provides a fairly detailed comparison of a number of image assessment criteria, including quantities based on the spot diagram, on o.t.f.'s and also on the variance of the wavefront aberration. The results do not provide an unequivocal answer as to whether, if the same initial optical system was optimised automatically, there would be a difference in the final system if different assessment functions were used. However, what is shown is that what differences do arise are likely to be small. Furthermore, it is shown that the value (if any) of full o.t.f. calculations during design is very dubious since other quantities which are far easier to calculate would allow the designer to rank systems similarly. The variance of the aberration difference function, in particular, is shown to be a very useful function since, by appropriate choice of spatial frequency, it may be used to agree with rankings of systems according to the radius of gyration of the spot diagram, or with the variance of the wavefront aberration or with the modulation transfer function (m.t.f.) when this is greater than 0.7.

To carry out this study it was decided that the aberration theory of Buchdahl provided a most useful tool. Most of the theory has been previously developed up to the seventh order terms of the transverse aberration expansion, and in some cases beyond this. Since the details for the derivatives of the seventh order coefficients had not been derived, nor those needed to include coordinate variations when dealing with higher order $\bar{O}T$ coefficients, these have been presented. To evaluate quantities associated with the wavefront aberration it was necessary to make use of the retardation coefficients. The writer is of the belief that almost no use has been made of them up until now. A brief convergence study is presented which confirms their potential usefulness, and this potential is realized by incorporating some modifications to the coefficients which greatly facilitate their use in expressions for the variance of the wavefront aberration and also that of the aberration difference function.

Use is also made of orthogonal expansions for the aberration functions to express the various image criteria as the sums of squares of orthogonal coefficients. The writer believes that these expressions present the most explicit formulation of the classical problem of aberration balancing, for optimization is then the process of reducing any of these coefficients towards zero. The coefficients are linear combinations of the classical aberration coefficients, the form of the combination being dependent on the particular function chosen. A number of authors have considered such orthogonal expansions for the wave aberration and the aberration difference function. However it would appear that this has not been done for the transverse aberration expansion, and that explicit formulation of the orthogonal coefficients in the other cases has not been carried out to the detail given here.

The results obtained from analytic comparison of rather simple cases have been complemented by considerable numerical comparisons, and the results have shown surprisingly good correlation granted the simplicity of the analytic cases examined.

Considerable reference is made throughout the work on aberration theory to publications of H.A. Buchdahl. To facilitate such reference it is convenient to use the following abbreviations:-

L: Optical Aberration Coefficients, Dover Publications Inc. New York (1968).

H: An Introduction to Hamiltonian Optics, Cambridge University Press, London (1960).

O.A.C.I: refers to the first paper of a series of papers on Optical Aberration Coefficients published in the Journal of the Optical Society of America. Others in the series are likewise denoted by the prefix OAC followed by the number of the paper.

Many symbols are used in the text and to help overcome some of the inevitable problems associated with following the theory a glossary of symbols is given at the end of the thesis, prior to the appendices.

The notation and terminology used largely follows Buchdahl(L) in those parts concerned with the Lagrangian aberration theory. The following points of notation are worth noting explicitly:-

a) Since we are often dealing with approximations to functions it is of value to know the order of approximation. This is denoted by the symbol O(n) where n is the degree of the terms of lowest order neglected.

b) Many quantities dealt with are vectors and to avoid duplication of equations and expressions a bar under a quantity denotes both y and z components of that quantity; thus $\underline{X}$ stands for (X_y, X_z).

c) Since many quantities are expressed as the sums of terms of different order it is convenient to be able to concisely distinguish such terms. Hence, if X is approximated by a series of terms of even order we write

$$X = \sum_{n=0}^{\infty} X^{(2n)}$$

where $X^{(2n)}$ contains only terms of degree 2n in the expansion variables.

ACKNOWLEDGEMENTS

I am much indebted to Dr. F.D. Cruickshank for introducing me to the theory of optical design and suggesting the study presented in this thesis. I thank him for his critical comments and willing help, and in particular for his continued efforts well beyond retirement from the University staff. The supervision and assistance of Dr. M.D. Waterworth over the latter stages of this work are also much appreciated.

In addition to the above I also wish to thank the following:-

Professor H.A. Buchdahl and Dr. P.J. Sands who willingly provided critical appraisal and valuable discussion of various aspects of this work;

the staff of the Hydro-University Computing Centre for the generous provision of computing facilities;

Mrs. B. Thomson, Mrs. R. Clothier and Mrs. J. McLean who typed the thesis, and Mrs. M. Woodruff and Mrs. H. Hen who prepared the diagrams;

the Australian Government for the financial assistance of a Commonwealth Post-Graduate Research award.

Finally I must express my deep gratitude to my parents and my wife who have given me continuing help and encouragement during the various stages of my studies.

CHAPTER 1

GEOMETRICAL OPTICS AND ABERRATION EXPANSIONS

§1.1 Introduction

Since the publication of Seidel's work on geometrical optics, power series representations of the departure of an optical system from paraxial behaviour have been much used in geometrical theory. The coefficients of these series have been approximated by the tracing of selected finite rays through the system and using the resultant discrete values of the aberration to determine the best-fitting polynomial to some aberration function[1]. This method is rapid and procedures to ensure its reliability have been developed[2]. Analytic methods for determining the coefficients have also been developed, most notably by Buchdahl[3]. The usefulness of these analytic methods in design depends on the rate of convergence of the power series concerned, and no comprehensive analysis of this is available. However, empirical investigations have established the approximate range of satisfactory convergence for different types of systems when various expansions, and terms of these expansions, are used[4].

Though the calculation of higher order analytic coefficients is not as simple a task as that of calculating coefficients in approximating polynomials, the availability of these analytic expressions has practical value through their usefulness in providing computationally efficient methods for calculating the derivatives, of any order, of the aberration with respect to constructional parameters. Also the representation of the coefficients as the sum of contributions from each surface allows the designer to examine in detail the part played by individual surfaces in detracting from stigmatic imagery. The superiority of the analytic method increases rapidly as the order of derivative increases. Also, since the coefficients calculated analytically are field-independent, the extent of spatial variance in image formation by a given system may be easily deduced.

Buchdahl has developed expressions for calculating the third, fifth, seventh and ninth order coefficients in the expansion for the transverse ray aberration of an optical system, and the fourth, sixth and eighth order coefficients in the expansion for the wavefront aberration[5]. Also methods have been given for the computation of the derivatives of the transverse ray aberration coefficients with respect to axial and extra-axial curvatures, surface separations and refractive indices, with details provided for the third and fifth order coefficients. In this chapter the above theory is summarised and the writer has extended it by giving details necessary for the extension of the derivatives to tertiary coefficients. In addition, some numerical results on the convergence of the series expansion of the wavefront aberration (or retardation) function are presented and the results show the expected agreement with results presented by Ford on the transverse ray aberration series. The relation between entrance pupil coefficients and the retardation coefficients is also discussed, with emphasis on the physical interpretation of the relationship, and finally some approximate relations between the two types of coefficients are given.

§1.2 Coordinate Systems and Nomenclature

Let S be a symmetric optical system. Let Oxyz be a set of left-handed rectangular cartesian coordinates with the x-axis lying along the optical axis of the system, and the x-y plane containing the object point. Let the origin 0 lie at the axial point of the first surface. Let c_{0j} denote the axial curvature of the j^{th} surface, d_j denote the axial distance from the $(j-1)^{th}$ surface to the j^{th} surface (with $d_1 \equiv 0$), and N_j

be the refractive index of the medium preceding surface j.

Any ray is specified by its coordinates (Y_1, Z_1) in the polar tangent plane of the first surface, that is, in the plane x=0, and its direction tangents (V_1, W_1) at that point. Thus any ray is specified by its four coordinates (Y_1, Z_1, V_1, W_1), conveniently denoted by $(\underline{Y}_1, \underline{V}_1)$. In the plane $x = \sum_{i=1}^{j} d_i$, i.e. the polar tangent plane of the j^{th} surface, the ray passes through the point $(\underline{Y}_j)$ with direction $(\underline{V}_j)$. After refraction at surface j the ray variables in this plane are $(\underline{Y}'_j, \underline{V}'_j)$. Upper case symbols refer to coordinates and variables of finite rays, while the corresponding lower case symbols denote paraxial quantities.

It is often convenient to specify a ray in terms of coordinates other than the canonical coordinates $(\underline{Y}_1, \underline{V}_1)$ introduced in the previous paragraph. In the course of this work two other sets of coordinates introduced by Buchdahl will be used. These are referred to as paracanonical coordinates and are denoted by $(\underline{S}, \underline{T})$. The first set of paracanonical coordinates used are the $\bar{O}T$ coordinates, defined by

$$\underline{S} = (1-p/\ell_{01})^{-1}\, \underline{Y}_1 - p.(1-p/\ell_{01})^{-1}\, \underline{V}_1$$
$$\underline{T} = -\underline{Y}_1/\ell_{01} + \underline{V}_1$$

where ℓ_{01} is the object plane position and p is the entrance pupil position. The $\bar{O}T$ coordinates are thus seen to be linearly related to the canonical coordinates. The second set of coordinates used are the W-coordinates which are a more general form of paracanonical coordinates. They are defined by

$$\underline{S} = e^{-1}\underline{Y}_{E'} = e^{-1}\underline{Y}'_k - p'e^{-1}\underline{V}'_k$$

$$\underline{T} = m\underline{H}_1 \quad = -m\underline{Y}_1 + m/\ell_{01}.\underline{V}_1$$

where e is the axial distance from the paraxial exit pupil to the ideal image plane, p′ is the position of the paraxial exit pupil, and m is the paraxial magnification of the system for a given object position ℓ_{01}.

§1.3 The Transverse Aberration

Optical design is a non-trivial problem because of the failure of optical systems to satisfy exactly the laws of paraxial optics. To analyse the departure from paraxial behaviour of rays traversing the system various aberration functions have been introduced. The transverse aberration function is one such quantity. It gives the vector displacement of the intersection point of some ray with the gaussian image plane from the corresponding paraxial ray intersection point. This displacement is a function of the coordinates of the ray. By choosing a suitable set of coordinates for specifying rays it is possible to expand the aberration function as a power series, the coefficients of which are known as aberration coefficients.

For a symmetric optical system the aberration function can be expressed in terms of three rotational invariants[6].

$$\xi = S_y^{\,2} + S_z^{\,2} \quad , \quad \eta = Sy\ Ty + Sz\ Tz \quad , \quad \zeta = T_y^{\,2} + T_z^{\,2} \qquad (1.3.1)$$

Using Buchdahl's notation the expansion can then be written

$$\underline{\varepsilon}'_k = \mu \sum_{n=1}^{\infty} \sum_{\mu=0}^{n} \sum_{\nu=0}^{\mu} \left(G^{(n)}_{\mu\nu k}\, \underline{S} + \bar{G}^{(n)}_{\mu\nu k}\, \underline{T} \right) \xi^{n-\mu}\, \eta^{\mu-\nu}\, \zeta^{\nu} \qquad (1.3.2)$$

where $\underline{\varepsilon}'_k$ is the aberration of the ray with coordinates $(\underline{S},\underline{T})$, $\mu = (N'_k \nu'_{ok})^{-1}$, the subscript o indicating that the ray with which that symbol is associated passes through the axial point of the object plane. The $G^{(n)}_{\mu\nu k}$, $\bar{G}^{(n)}_{\mu\nu k}$ are aberration coefficients of order 2n+1.

Details of a method for the calculation of the aberration coefficients have been given by Buchdahl[3]. This method is based on the quasi-invariance of the quantity

$$\underline{\Lambda}_j = N_j\, v_{oj}\, \underline{H}_j \qquad (1.3.3)$$

for a given ray, where

$$\underline{H}_j = \ell_{oj}.\underline{V}_j - \underline{Y}_j \quad , \quad \ell_{oj} = y_{oj}/v_{oj} \tag{1.3.4}$$

Each iteration gives the exact aberration coefficients for successively higher orders. It is of interest to note that when coordinates which are not linearly related to the canonical coordinates are used it may be necessary to carry out the iterations surface-by-surface rather than order-by-order. Sands[7] has discussed this situation in some detail. The W-coordinates are an example of a set of coordinates which are not linearly related to the canonical coordinates, but which may still be found using an order-by-order process of iteration. Buchdahl[8] has also given a transformation method for calculating the W-coefficients from some other set of paracanonical coefficients. This is the method used for this study.

§1.4 The Retardation Function

Rather than deal with the transverse aberration of a ray it is often convenient to consider another geometrical quantity known as the wavefront aberration or the retardation of the wavefront. The retardation is defined as follows:-

Let P´ be the gaussian image of a point P formed by a system S with exit pupil plane E'. Let E´ be the axial point of E'.

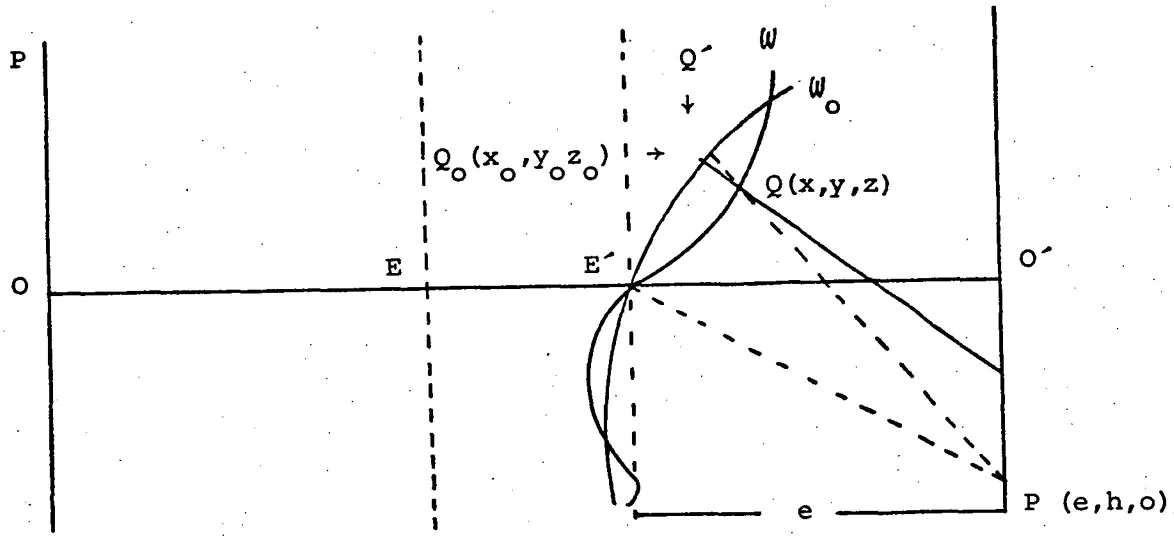


fig. 1.1

Let W_o be a sphere centred on P´ and passing through E´, and let W be a continuous set of points, containing E´, which are at equal optical distances from P. Let Q(x,y,z) be any point on W and $Q_o(x_o,y_o,z_o)$ the point in which the normal to W at Q intersects W_o. Then the retardation is

$$R(y,z,h) = Q_oQ$$

This definition is in agreement with that of Born and Wolf[9], and Buchdahl in H[10], but differs from that given in reference 5 and by Rayces[11]. A discussion of these differences is included in appendix 1.

The retardation function R may be expanded in a power series of the form

$$R(y,z,h) = \sum_{n=2}^{\infty} \sum_{i=0}^{n} \sum_{j=0}^{i} w_{ij}^{(n)} \lambda^{n-i} \mu^{i-j} \nu^{j} \tag{1.4.1}$$

where $\lambda = y^2+z^2, \mu = yh, \nu = h^2$, h being the paraxial image height and (y,z) coordinates of a ray on the actual wavefront. The $w_{ij}^{(n)}$ are retardation coefficients of order 2n.

Expressions for the retardation coefficients as functions of the constructional parameters of the system are obtained by establishing exact relationships between these coefficients and the W coefficients via the deformation function. The deformation function is defined thus:-

Referring to fig. 1.1, let Q' lie on W_o such that $P'QQ'$ are collinear. Then the deformation function is

$$D(y,z,h) = \tfrac{1}{2}e^{-1}[(Q'P')^2 - (QP')^2] \tag{1.4.2}$$

The significance of the deformation function lies in the fact that the exact relationship

$$\underline{\varepsilon} = e\frac{\partial D}{\partial \underline{y}} \tag{1.4.3}$$

holds. If $\underline{\varepsilon}$ is expressed as a function of the coordinates (y,z,h) and both $\underline{\varepsilon}$ and D are expanded as power series then relations between the ray aberration coefficients and the deformation coefficients may be found.

Now $QQ_o = QQ' + O(10)$ (1.4.4)

and $R = (1-\tfrac{1}{2}e^{-2}\nu + \tfrac{3}{8}e^{-4}\nu^2)D + \tfrac{1}{2}e^{-1}D^2 + O(10)$ (1.4.5)

so exact relations between the fourth, sixth and eighth order retardation and deformation coefficients can be found. Thus, via the deformation function, exact relations between the retardation coefficients and the transverse aberration coefficients may be found.

§1.5 Convergence of the Retardation Function Series

As pointed out earlier, no analytic method for determining the convergence of the aberration expansions is available. However for transverse aberrations comparisons between the aberrations found by ray trace and those predicted using coefficients have been given for $\bar{O}T$ coefficients by Buchdahl[12] and Ford[4], and for off-axis coefficients by Sands[13]. Since nowhere in the literature does there appear any mention of the use of the retardation coefficients, let alone examination of the convergence of the retardation expansion, a brief comparison of raytraced and predicted wavefronts is presented here.

Ford's raytrace procedure was modified to allow storage of inter-surface optical distances for a principal ray and the differences between these and the corresponding distances for other rays were found and summed to determine the retardation. Significant round-off errors occurred initially and a modification to eliminate this is given in appendix 2.

In the following tables the differences, δ, between the wavefront aberration found by raytrace and that using the fourth, sixth and eighth order terms of the aberration expansion are given. The unit chosen corresponds to the D line for systems of 1 inch focal length. The systems examined are photographic objectives of various types - four given in NBS monograph 93 by Stavroudis and Sutton, four used by Ford in his examination of the convergence of the transverse aberration expansion and one being an original design by the author using Cruickshank's method (see chapter 4). For each system a set of angles is chosen and at each angle the agreement between predicted and raytraced aberration is examined at 0.5, 0.75 and 1.0 times the aperture for which the system is designed to be used. The tables show the value of δ, the value of δ as a percentage of the maximum aberration in either section, and a subjective assessment (G = good, F = fair, P = poor) of the quality of agreement. The wavefront is examined in both sagittal (S)

	Field angle		0°			5°			10°			15°		
	Fraction of full aperture		0.5	0.75	1.0	0.5	0.75	1.0	0.5	0.75	1.0	0.5	0.75	1.0
Sagittal section	actual error (in units of λ)	δ	0	0.01	1.2	0	0.5	1.6	0	0.1	1.8	0.05	0.1	1.6
	percentage error	%	0	0.2	9	0	7	12	0	1	18	0.06	1.0	16
	subjective quality	Q	G	G	F	G	F	F	G	G	F	G	G	F
Tangential section		δ	0	0.01	1.2	0	0.8	3	0	1.5	6	0.7	3.2	Vig
		%	0	0.02	9	0	9	20	0	15	60	8	32	Vig
		Q	G	G	F	G	F	F	G	F	P	F	P	-

Table 1.1

NBS system (p. 36 of NBS monograph 93). F/2.5

	0°			2.5°			5.0°			7.5°			10.0°			12.5°		
	0.5	0.75	1.0	0.5	0.75	1.0	0.5	0.75	1.0	0.5	0.75	1.0	0.5	0.75	1.0	0.5	0.75	1.0
δ	0	0	0.04	0	0	0.03	0	0	0.03	0	0	0.03	0	0.01	0.03	0	0	0.02
%	0	0	2	0	0	1.5	0	0	1.4	0	0	2	0	1	2	0	0	2
Q	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G
δ	0	0	0.04	0	0	0.04	0	0.01	0.05	0	0.01	0.06	0	0.02	0.05	0	0.02	0.05
%	0	0	2	0	0	2	0	0.6	2.3	0	1	4	0	2	3	0	2	5
Q	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	F

Table 1.2

NBS system (p. 46). F/3.5

		0°			2.5°			5.0°			7.5°			10.0°		
		0.5	0.75	1.0	0.5	0.75	1.0	0.5	0.75	1.0	0.5	0.75	1.0	0.5	0.75	1.0
Sagittal section	δ % Q	0 0 G	0 0 G	0.03 1 G	0 0 G	0 0 G	0.03 1 G	0 0 G	0 0 G	0.03 1 G	0.01 0.5 G	0 0 G	0.03 0.6 G	0 0 G	0 0.1 G	0.03 0.7 G
Tangential section	δ % Q	0 0 G	0 0 G	0.03 1 G	0 0 G	0 0 G	0.03 1 G	0 0 G	0.01 0.4 G	0.04 1 G	0 0.3 G	0.01 0.5 G	0.04 0.8 G	0 0 G	0.01 0.2 G	0.05 1.2 G

Table 1.3

NBS system (p. 56). F/6.3

		0°			2°			4°			6°			8°		
		0.5	0.75	1.0	0.5	0.75	1.0	0.5	0.75	1.0	0.5	0.75	1.0	0.5	0.75	1.0
	δ % Q	0 0 G	0.01 0.3 G	0.22 5 F	0 0 G	0.02 0.7 G	0.18 4 G	0 0 G	0.02 0.8 G	0.26 0.5 G	0 0 G	0.01 0.2 G	0.22 3 G	0 0 G	0.04 1 G	0.25 3 G
	δ % Q	0 0 G	0.01 0.3 G	0.22 5 F	0 0 G	0.03 1 G	0.24 5 F	0 0 G	0.01 0.4 G	0.25 0.5 G	0 0 G	0.08 2 G	Vig Vig F	0.02 0.5 G	0.12 3 G	Vig Vig -

Table 1.4

NBS system (p. 71). F/6

	0°			2°			4°			6°			8°			10°		
	0.5	0.75	1.0	0.5	0.75	1.0	0.5	0.75	1.0	0.5	0.75	1.0	0.5	0.75	1.0	0.5	0.75	1.0
δ	0	0	0.04	0	0.01	0.04	0	0	0.04	0	0	0.03	0	0.01	0.04	0	0	0
%	0	0	1.1	0	0.5	1.2	0	0	1	0	0	0.6	0.3	0.2	0.6	0	0	0.4
Q	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G
δ	0	0	0.04	0	0.01	0.05	0	0	0.07	0	0.01	0.07	0.01	0.04	0.07	0.06	0.21	0.6
%	0	0	1.1	0	0.6	1.4	0	0.1	1.7	0	0.8	1.4	0.6	1.3	1	4	4	7
Q	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	F

Table 1.5

Original triplet design by author.

	0°			10°			20°			30°			35°			40°		
	0.5	0.75	1.0	0.5	0.75	1.0	0.5	0.75	1.0	0.5	0.75	1.0	0.5	0.75	1.0	0.5	0.75	1.0
δ	0	0	0.05	0	0	0.02	0	0.01	0.03	0.08	0.2	0.32	0.3	0.8	1.2	1.3	2.3	4.5
%	0	0	2.5	0	0	2	0	1	2	20	18	12	150	100	55	>200	>200	170
Q	G	G	G	G	G	G	G	G	G	F	F	F	P	P	P	P	P	P
	0	0	0.05	0	0	0.01	0	0.01	Vig	0.06	0.04	0.08	0.05	0.1	0.2	0.1	0.15	0.3
%	0	0	2.5	0	0	1	1	1	Vig	15	3	3	22	12	9	40	17	11
Q	G	G	G	G	G	G	G	G	–	F	G	G	F	F	F	P	F	F

Table 1.6

Tropogon (Ford). F/6.3

	0°			5°			10°			15°			20°			25°		
	0.5	0.75	1.0	0.5	0.75	1.0	0.5	0.75	1.0	0.5	0.75	1.0	0.5	0.75	1.0	0.5	0.75	1.0
δ	0	0.07	0.9	0	0	0.2	0	0.06	0.33	0.06	0.12	0.4	0.4	0.8	1.6	0.9	2.7	5.2
%	0	1	8	0	0	2	0	0.8	2.7	1.1	1.0	2.5	7	8	14	30	110	35
Q	G	G	F	G	G	G	G	G	G	G	G	G	F	F	F	P	P	P
δ	0	0.07	0.9	0	0	0.33	0	0.02	0.18	0.11	0.4	2	1.8	4.2	16	2.1	4.8	Vig
%	0	1	8	0	0	3	0	0.3	1.4	2	3.7	12	28	42	160	68	200	Vig
Q	G	G	F	G	G	G	G	G	G	G	G	F	F	P	P	P	P	P

Table 1.7

Sonnar (Ford). F/1.6

	0°			2°			4°			6°			8°			10°		
	0.5	0.75	1.0	0.5	0.75	1.0	0.5	0.75	1.0	0.5	0.75	1.0	0.5	0.75	1.0	0.5	0.75	1.0
δ	0	0	0.05	0	0	0.05	0	0	0.01	0	0	0.02	0	0	0.01	0	0	0.02
%	4	2	16	2	0	11	0.4	0.3	0.6	0	0	0.5	0	0	0.2	0	0	1
Q	G	G	F	G	G	F	G	G	G	G	G	G	G	G	G	G	G	G
δ	0	0	0.05	0	0	0.04	0	0	0.03	0	0	0.05	0	0.01	0.04	0.01	0.04	0.11
%	4	2	16	2	1	9	0.4	0.2	1.8	0	0	1.1	0	0.4	0.8	0.6	0.9	1.3
Q	G	G	F	G	G	F	G	G	G	G	G	G	G	G	G	G	G	G

Table 1.8

Petzval (Ford). F/3.3

	0°			3°			6°			9°			12°			15°		
	0.5	0.75	1.0	0.5	0.75	1.0	0.5	0.75	1.0	0.5	0.75	1.0	0.5	0.75	1.0	0.5	0.75	1.0
δ	0.05	0.08	0.44	0.06	0.1	0.5	0.04	0.10	0.4	0.03	0.07	0.35	0.02	0.07	0.35	0.02	0.07	0.2
%	4.5	1.8	3.9	5	2.1	0.5	2.7	2.2	4.1	1.5	1.4	1.8	0.8	1.2	1.2	0.8	0.9	0.7
Q	G	G	G	G	G	G	G	G	G	G	G	F	G	G	G	G Q	G	G
δ	0.05	0.08	0.44	0.05	0.07	0.9	0.02	0.04	1.9	0.03	0.09	2.7	0.06	0.4	2.5	0.18	1.0	3.5
%	4.5	1.8	3.9	4	1.5	0.8	1.3	0.8	19	1.5	1.8	13	2.2	6.9	9	7.2	18	12
Q	G	G	G	G	G	G	G	G	P	G	G	P	G	F	F	F	F	P

Table 1.9

Biotar (Ford). F/1.6

and tangential (T) sections. For tangential section results the average is taken of upper and lower parts of the pupil, both for specifying the values of δ and for the maximum retardation for determination of percentage deviations.

The results here indicate that for small apertures (up to about F/5) and small semi-field angles (up to about 10°) the wavefront retardation expansion may be truncated after the eighth order terms with negligible loss in accuracy. For wider aperture systems the accuracy of the truncated expansion is not as reliable, as indicated by the results of tables 1.1 and 1.7. However an examination of the degree of vignetting in both these systems indicates that the results are not as serious as might be thought, since in both cases (particularly for the Sonnar) vignetting would eliminate are large proportion of those rays for which the retardation expansion is inaccurate. The results of tables 1.2, 1.3 and 1.9 indicate that quite good convergence may be obtained for wider aperture systems, however, and if the aperture is small larger field angles may be covered, as indicated by the results for the Topogon (table 1.6).

The results obtained using the systems studied by Ford indicate the expected substantial correlation in accuracy of the corresponding truncated expansions; for those cases where the transverse aberration series loses accuracy, however, it would appear that the wavefront retardation series is even less reliable. This is, perhaps, understandable in view of the relation (1.4.3), showing that the wavefront aberration is related to the integral of the ray aberration and hence errors in the latter are "summed" to give errors in the former.

As a useful (but nof foolproof) general rule it would seem that one can assume that for systems designed to work at 20° semi-field and F/1.6 the accuracy of the truncated expansion would be better than 10% for all but the extreme values of aperture and field. Beyond the region of apertures and field angles for which a particular system is designed not only does the aberration increase but also the inaccuracy of the aberration expansion generally increases dramatically. Thus a system designed to work at F/5 over a 10° semi-field will have an accurate aberration expansion for the whole of its aperture and field, but outside these limits the aberration expansion will normally become unreliable very rapidly.

§1.6 Differentiation of the Aberration Coefficients

a) Introduction: The systematic improvement of the imaging performance of an optical system requires a knowledge of the effect of the variation of a certain design parameter on the chosen merit function. Since the merit functions involve either the ray aberration or the wavefront retardation this means that the effect of varying a given parameter on the geometrical aberration needs to be found. If the aberration associated with a given ray is known as a polynomial function of the coordinates of that ray

$$N'_k v'_{ak} \underline{\varepsilon}'_k = \sum_{n=1}^{\infty} \sum_{\mu=0}^{n} \sum_{\nu=0}^{\mu} (G'^{(n)}_{\mu\nu k}\ \underline{S} + \bar{G}'^{(n)}_{\mu\nu k}\ \underline{T})\,\xi^{n-\mu}.\eta^{\mu-\nu}.\zeta^{\nu} \qquad (1.6.1)$$

where $\underline{S}$ and $\underline{T}$ are coordinates of the ray

$$\xi = S_y^{\ 2} + S_z^{\ 2}, \qquad \eta = S_y T_y + S_z T_z, \qquad \zeta = T_y^{\ 2} + T_z^{\ 2}$$

then
$$\frac{\partial \underline{\varepsilon}'_k}{\partial \tau_j} = \frac{1}{N'_k v'_{ak}} \sum_{n=1}^{\infty} \sum_{\mu=0}^{n} \sum_{\nu=0}^{\mu} (F^{(n)}_{\mu\nu k}\ \underline{S} + \bar{F}^{(n)}\ \underline{T})\,\xi^{n-\mu}.\eta^{\mu-\nu}\zeta^{\nu} \qquad (1.6.2)$$

where
$$F^{(n)}_{\mu\nu k} = \frac{G^{(n)}_{\mu\nu k}}{\partial \tau_j} - \frac{v'_{ak}}{\partial \tau_j}\cdot\frac{G'^{(n)}_{\mu\nu k}}{v'_{ak}}$$

τ_j being the parameter which is varied.

Thus to find the effect of a variation of a parameter on the image assessment function it is necessary to calculate the derivatives of the augmented aberration coefficients with respect to that parameter.

Suppose the merit function is some function of the ray distribution in the gaussian image plane, such as the radius of gyration, k, where

$$k = A^{-1}.\iint_A \varepsilon^2(\rho,\theta)\,dA \tag{1.6.3}$$

$\varepsilon(\rho,\theta)$ is the aberration of the ray having polar coordinates (ρ,θ) in the equivalent entrance pupil and A is the equivalent entrance pupil. Then

$$k = f(G_{\mu\nu}^{(n)},\ \bar{G}_{\mu\nu}^{(n)},\ \rho_o) \tag{1.6.4}$$

assuming a circular entrance pupil of radius ρ_o.

In using (1.6.~~3~~4) above the aberration coefficients have a certain geometrical significance and this must be retained when the change $\delta\tau_j$ is made. The $\bar{O}$T-coordinates, which will be used in obtaining spot diagram based assessment quantities in the following work, are dependent on the pupil position and so the effect on this of varying τ_j must be found.

In the course of this work details necessary for the calculation of derivatives with respect to axial curvatures are given. However where the expressions apply for any system parameter this is indicated by the use of the symbol τ_j rather than c_{oj} to specify the parameter.

b) <u>Differentiation of $\bar{O}$T-coordinates with respect to axial curvatures</u>

The aberration of a ray can be expressed as a power series in the ray coordinates as in (1.4.1). If the parameter τ_j is varied by $\delta\tau_j$ then the coefficients of this series are changed. It is the derivatives of these coefficients which are now required. The method of calculation is given in L part III, and is outlined here with particular reference to variation of the axial curvatures.

Let δc_{oj} be a small change in c_{oj}. This leads to a change in the optical properties of surface j only. Now

$$N_k' v_{ak}' \underline{\varepsilon}_k' = \underline{\hat{\varepsilon}}_k' = \sum_{j=1}^{k} \Delta\underline{\Lambda}_j \tag{1.6.1}$$

so $\Delta\underline{\Lambda}_j$ is changed in form when c_{oj} is changed - i.e. the intrinsic coefficients are changed. The quantities $\Delta\underline{\Lambda}_i$ $(i<j)$ are unaltered, while for $i>j$ the $\Delta\underline{\Lambda}_i$ are altered, not in form, but through the resultant change in the paraxial and non-paraxial variables occurring in the expressions for $\Delta\underline{\Lambda}_i$ $(i>j)$. Now

$$\underline{Y}_j' = y_{aj}.(\underline{S} + \underline{\delta}_{sj}) + y_{bj}.(\underline{T} + \underline{\delta}_{tj}) \tag{1.6.2}$$

$$\underline{V}_j' = v_{aj}'.(\underline{S} + \underline{\delta}_{sj}) + v_{bj}'.(\underline{T} + \underline{\delta}_{bj})$$

where $\underline{\delta}_{sj} = -\sum_{i=1}^{j} \Delta\underline{\Lambda}_{bi}$, $\underline{\delta}_{bj} = \sum_{i=1}^{j} \Delta\underline{\Lambda}_{ai}$. By rearranging the equations

(1.6.2) and considering a variation in c_{oj} it follows that

$$\partial\underline{S} = `N_j'.(v_{bj}'.\partial\underline{Y}_j' - y_{bj}.\partial\underline{V}_j') + \partial\, `\underline{\hat{\varepsilon}}_{bj}'$$

$$\partial\underline{T} = `N_j'.(-v_{aj}'.\partial\underline{Y}_j' + y_{aj}.\partial\underline{V}_j') - \partial`\underline{\hat{\varepsilon}}_{aj}' \tag{1.6.3}$$

From (1.6.2)

$$\partial\underline{Y}_j' = y_{aj}.\partial(\underline{\delta}_{sj}') + y_{bj}.\partial(\underline{\delta}_{tj}') = (\partial.\Delta\underline{\Lambda}_j|y_j) \tag{1.6.4a}$$

$$\partial\underline{V}_j' = v_{aj}'.\partial(\underline{\delta}_{sj}') + v_{bj}'.\partial(\underline{\delta}_{tj}') + (1-k_j)\underline{Y}_j'.\delta c_{oj}$$

so $$\partial\underline{V}_j' = (\delta\Delta`\underline{\Lambda}_j|v_j') + (1-k_j).\underline{Y}_j'.\partial c_{oj} \tag{1.6.4b}$$

where the notation $(x|y) = x_a y_b - x_b y_a$ has been used.

Substituting for $\partial\underline{Y}_j', \partial\underline{V}_j'$ in (1.6.3) using (1.6.4) gives

$$\frac{\partial\underline{S}}{\partial c_{oj}} = -(N_j'-N_j).y_{bj}\underline{Y}_j' + \frac{\partial`\underline{\varepsilon}'_{bj}}{\partial\chi}.\frac{\partial\chi}{\partial c_{oj}} - \frac{\partial\Delta\underline{\Lambda}_{bj}}{\partial c_{oj}}$$

$$\frac{\partial \underline{T}}{\partial c_{oj}} = (N_j'-N_j).y_{aj}\underline{Y}_j' - \frac{\partial'\hat{\underline{\varepsilon}}_{aj}}{\partial\chi}.\frac{\partial\chi}{\partial c_{oj}} + \frac{\partial\Delta'\underline{\Lambda}_{aj}}{\partial c_{oj}} \qquad (1.6.5)$$

where χ indicates $\partial Y_1, \partial V_1, \partial Z_1, \partial W_1$ successively.

Now since the effect of varying c_{oj} on $\Delta\underline{\Lambda}_i$, $i>j$, is simply to change the values of the paraxial and non-paraxial variables, this can be considered as equivalent to a change in the coordinates of the ray. The problem then is to determine this equivalent variation in coordinates. So

$$\frac{\partial\hat{\underline{\varepsilon}}_k'}{\partial c_{oj}} = \frac{\partial(\hat{\underline{\varepsilon}}_k'-\hat{\underline{\varepsilon}}_j')}{\partial y_{ol}}.\frac{\partial y_{ol}}{\partial c_{oj}} + \frac{\partial(\hat{\underline{\varepsilon}}_k'-\hat{\underline{\varepsilon}}_j')}{\partial v_{ol}}.\frac{\partial v_{ol}}{\partial c_{oj}} + \frac{\partial(\hat{\underline{\varepsilon}}_k'-\hat{\underline{\varepsilon}}_j')}{\partial\chi}.\frac{\partial\chi}{\partial c_{oj}} + \frac{\partial\Delta\underline{\Lambda}_j}{\partial c_{oj}}, \qquad (1.6.6)$$

since $\sum_{i=j+1}^{k} \Delta\underline{\Lambda}_i = \hat{\underline{\varepsilon}}_k'-\hat{\underline{\varepsilon}}_j'$. The equivalent variation of coordinates will be of the form

$$\partial\underline{S} = (\Gamma_a\underline{S} + \bar{\Gamma}_a\underline{T}).\partial c_{oj}$$

$$\partial\underline{T} = (\Gamma_b\underline{S} + \bar{\Gamma}_b\underline{T}).\partial c_{oj} \qquad (1.6.7)$$

where $\Gamma_a = \sum_{n=0}^{\infty}\sum_{\mu=0}^{n}\sum_{\nu=0}^{\mu} \gamma_{\mu\nu a}\xi^{n-\mu}.\eta^{\mu-\nu}.\zeta^{\nu}$ and likewise for $\bar{\Gamma}_a, \Gamma_b, \bar{\Gamma}_b$.

Only the zero, second and fourth order terms in these expansions are needed to calculate derivatives of paraxial, third, fifth and seventh order coefficients. Hence it is convenient to write

$$\Gamma_a = \Theta_a + (\alpha_a\xi + \beta_a\eta + \gamma_a\zeta) + \sigma_{1a}\xi^2 + \sigma_{2a}\xi\eta + \ldots + \sigma_{6a}\zeta^2 + \ldots. \qquad (1.6.8)$$

and, again, similarly for $\bar{\Gamma}_a, \Gamma_b, \bar{\Gamma}_b$.

Now $\hat{\underline{\varepsilon}} = A\underline{S}\xi + \bar{A}\underline{T}\xi + B\underline{S}\eta + \bar{B}\underline{T}\eta + .. + S_1\underline{S}\xi^2 + \bar{S}_1\underline{T}\xi^2 + ..$

and so

$$\frac{\partial\hat{\varepsilon}_y'}{\partial\chi}.\frac{\partial\chi}{\partial c_{oj}} = [\Gamma_a(S_y.\frac{\partial}{\partial S_y} + 2\xi.\frac{\partial}{\partial\xi} + \frac{\eta\partial}{\partial\eta}) + \bar{\Gamma}_a.(T_y\frac{\partial}{\partial S_y} + \frac{2\eta\partial}{\partial\xi} + \frac{\zeta\partial}{\partial\eta})$$

$$+ \Gamma_b.(S_y.\frac{\partial}{\partial T_y} + \frac{\xi\partial}{\partial\eta} + \frac{2\eta\partial}{\partial\zeta}) + \bar{\Gamma}_b.(T_y\frac{\partial}{\partial T_y} + \frac{\eta\partial}{\partial\eta} + \frac{2\zeta\partial}{\partial\zeta})].\varepsilon_y \qquad (1.6.9$$

Replacing $\Gamma_a, \bar{\Gamma}_a, \Gamma_b, \bar{\Gamma}_b$ by their expansions then gives $\frac{\partial \hat{\varepsilon}}{\partial \chi} \cdot \frac{\partial \chi}{\partial c_{oj}}$ as a power series in ξ, η, ζ as required. The primary and secondary expressions are given in L (121.6), while the tertiary expressions are given in appendix 3.

It still remains to obtain the derivatives of the j^{th} surface contribution to the aberration, that is $\frac{\partial(\Delta\underline{\Lambda}_j)}{\partial c_{oj}}$. This is most simply done by differentiating explicitly the expressions for the intrinsic coefficients and then the expressions for the total surface contributions. The latter simply involves the differentiation of the expressions L (11.3), (81.3). To obtain the derivatives of the intrinsic coefficients one may use the methods outlined by Buchdahl which are computationally efficient for the secondary coefficients. However when tertiary coefficients and their derivatives are being considered a more direct approach through direct differentiation of expressions in L §§77,78 is straightforward and reasonably efficient. From L (60.3), if, for example, τ corresponds to c_o,

$$\Delta\underline{\Lambda} = \mathcal{D}\underline{I} + \mathcal{L}\underline{Y} \tag{1.6.10}$$

so
$$\frac{\partial \Delta\underline{\Lambda}}{\partial \tau_j} = \underline{I} \cdot \frac{\partial \mathcal{D} \cancel{\underline{I}}}{\partial \tau_j} + \left(\mathcal{D} \cdot \frac{\partial c_{oj}}{\partial \tau_j} + \frac{\partial \mathcal{L}}{\partial \tau_j}\right)\underline{Y} \tag{1.6.11}$$

Now $\mathcal{D} = \sum_{i=1}^{\infty} \mathcal{D}^{(i)}$, $\mathcal{L} = \sum_{i=1}^{\infty}$ where $\mathcal{D}^{(i)}$ and $\mathcal{L}^{(i)}$ are polynomials of degree i in ξ, η and ζ, so

$$\frac{\partial \Delta\underline{\Lambda}^{(n)}}{\partial \tau_j} = \frac{\partial \mathcal{D}^{(n)}}{\partial \tau_j}\underline{I} + \left(\mathcal{D}^{(n)} \cdot \frac{\partial c_{oj}}{\partial \tau_j} + \frac{\partial \mathcal{L}^{(n)}}{\partial c_{oj}}\right)\underline{Y} \tag{1.6.12}$$

Using the expressions L (77.1-5), (78.1-4) the terms $\mathcal{D}_j^{(n)}$ and $\mathcal{L}_j^{(n)}$ are given in terms of ξ_j, η_j and ζ_j. The intrinsic coefficients are obtained from these by substituting for ξ_j, η_j and ζ_j using the paraxial approximations

$$\begin{aligned}
\xi_j &= y_{aj}^2 \xi_1 + 2y_{aj}y_{bj}\eta_1 + y_{bj}^2 \zeta_1 \\
\eta_j &= y_{aj}v_{aj}\xi_1 + (y_{aj}v_{bj} + y_{bj}v_{aj}) \cdot \eta_1 + y_{bj}v_{bj}\zeta_1 \\
\zeta_j &= v_{aj}^2 \xi_1 + 2v_{aj}v_{bj}\eta_1 + v_{bj}^2 \zeta_1
\end{aligned} \tag{1.6.13}$$

and $\underline{I}_j = i_{aj}\underline{S} + i_{bj}\underline{T}$, $\quad \underline{Y}_j = y_{aj}\underline{S} + y_{bj}\underline{T}$

To obtain $\dfrac{\partial \mathcal{D}^{(n)}}{\partial \tau_j}$ it is necessary to differentiate the expression for

$\mathcal{D}_j^{(n)}$. Let $\mathcal{D}_j^{(n)} = \sum\limits_{\mu=0}^{n} \sum\limits_{\nu=0}^{\mu} d_{\mu\nu j}^{(n)} \xi_j^{\,n-\mu} . \eta_j^{\,\mu-\nu} . \zeta_j^{\,\nu}$

$$= \sum_{\mu=0}^{n} \sum_{\nu=0}^{\nu} \underset{\sim}{g}_{\mu\nu j}^{(n)} \xi_1^{\,n-\mu} . \eta_1^{\,\mu-\nu} . \zeta_1^{\,\nu} \qquad (1.6.14)$$

Then the $\underset{\sim}{g}_{\mu\nu j}$ are the intrinsic coefficients of surface j.

$$\frac{\partial \mathcal{D}_j^{(n)}}{\partial \tau_j} = \sum_{\mu=0}^{n} \sum_{\nu=0}^{\mu} \frac{\partial d_{\mu\nu j}}{\partial \tau j} . \xi_j^{\,n-\mu} . \eta_j^{\,n-\nu} . \zeta_j^{\,\nu}$$

since $\dfrac{\partial \xi_j}{\partial \tau_j} = \dfrac{\partial \eta_j}{\partial \tau_j} = \dfrac{\partial \zeta_j}{\partial \tau_j} = 0$, provided $\tau_j \not\equiv N_j$. Note should be taken of the fact that variation of N_j affects the intrinsic properties of both surface j and surface j-1. A similar treatment can be applied to $L_j^{(n)}$ which is non-zero only in the presence of aspherics.

Finally it is necessary to consider the contribution to the derivative of the equivalent coordinate variations. Let

$\hat{\underline{\varepsilon}}_k' - \hat{\underline{\varepsilon}}_j' = \hat{\underline{\varepsilon}}_{[kj]}'$, so that

$$\frac{\partial \underline{\varepsilon}_k}{\partial c_{oj}} = \frac{\partial \underline{\varepsilon}_{[kj]}}{\partial s_{ol}} \cdot \frac{\partial s_{ol}}{\partial c_{oj}} + \frac{\partial \underline{\varepsilon}_{[kj]}}{\partial t_{ol}} \cdot \frac{\partial t_{ol}}{\partial c_{oj}} + \frac{\partial \underline{\varepsilon}_{[kj]}}{\partial \chi} \cdot \frac{\partial \chi}{\partial c_{oj}} + \frac{\partial \Delta \underline{\Lambda}_j}{\partial c_{oj}}, \qquad (1.6.15)$$

Now $\quad \dfrac{\partial s_1}{\partial c_{oj}} = -\bar{N}_j y_{bj} y_j$, $\quad \dfrac{\partial t_1}{\partial c_{oj}} = \bar{N}_j y_{aj} y_j \quad$ where $\bar{N} = N_j' - N_j$

so that, since $\hat{\underline{\varepsilon}}_k' = \hat{\underline{\varepsilon}}_{ak}' . s_{ol} + \hat{\underline{\varepsilon}}_{bk}' . t_{ol}$, (1.6.15) may be written

$$\frac{\partial \hat{\underline{\varepsilon}}_k'}{\partial c_{oj}} = \bar{N}_j . y_{aj} (y_j | \hat{\underline{\varepsilon}}_{[kj]}) + \frac{\partial \hat{\underline{\varepsilon}}_{[kj]}}{\partial \chi} \cdot \frac{\partial \chi}{\partial c_{oj}} + \frac{\partial \Delta \underline{\Lambda}_j}{\partial c_{oj}} \qquad (1.6.16)$$

Use of (1.6.16) gives the derivatives of the augmented aberration coefficients.

c) Differentiation with respect to pupil position

As mentioned earlier, when paracanonical coordinates are being used the effect of a change in the parameter τ_j may be to change the coordinate values for a particular ray. Hence the effect of this on the overall coefficients must be considered. Both $\bar{O}T$ and W coefficients depend on pupil position and so the coordinates may be affected by constructional parameter variations.

If the parameter which is varied precedes the diaphragm then there is a change δp in the location of the entrance pupil, and if the varied parameter follows the diaphragm then there is a change $\delta p'$ in the location of the exit pupil. Let the diaphragm be placed a distance b beyond the ℓ^{th} surface; then

$$p = (y_{a\ell}-v'_{q\ell}.b)/(v'_{p\ell}.b-y_{p\ell}) \qquad (1.6.17)$$

Let $p = -z_q/z_p$ where $z = y_\ell - v_\ell.b.$

Then $$\frac{\partial p}{\partial c_{oj}} = -(z_p.\frac{\partial z_q}{\partial c_{oj}} - z_q.\frac{\partial z_p}{\partial c_{oj}})/z_p^2 \quad \text{if } \ell<j$$

Now $$\frac{\partial z}{\partial c_{oj}} = \frac{\partial y_\ell}{\partial c_{oj}} - b.\frac{\partial v'_\ell}{\partial c_{oj}}$$

so $$\frac{\partial z}{\partial c_{oj}} = \bar{N}.(y_1|y_\ell)y_j - b.\bar{N}(y_j|v'_\ell).y_j$$

$$= \bar{N}(y_j|z).y_j$$

Hence $$\frac{\partial p}{\partial c_{oj}} = \bar{N}_p\{(p.y_{pj} + y_{qj}).(y_j|z)/z_q\}$$

i.e. $$\frac{\partial p}{\partial c_{oj}} = \bar{N}.(y_{pj}.p + y_{qj})^2 \qquad (1.6.18)$$

Suppose that $\ell \geq j$ now, so that $\delta p = 0$. Then $\delta p' \neq 0$.

$$p' = (y_{pk} \cdot p + y_{qk})/(v'_{pk} \cdot p + v'_{qk}) = z_{yk}/z_{vk}' \qquad (1.6.19)$$

$$\frac{\partial p'}{\partial c_{oj}} = p'/z_{yk}\left(\frac{\partial z_{yk}}{\partial c_{oj}} - p' \cdot \frac{\partial z_{vk}'}{\partial c_{oj}}\right)$$

$$= p' \cdot \bar{N} \cdot (y_{pj} \cdot p + y_{qj})/z_{yk}\{(y_j|y_k) - p' \cdot (y_j|v'_k)\}$$

so
$$\frac{\partial p'}{\partial c_{oj}} = \bar{N} p' z_{yj} \cdot (y_j | y_{E'})/z_{yk} \qquad (1.6.20)$$

where $y_{E'} = y_k - p' \cdot v'_k$

Equations (1.6.18) and (1.6.20) give expressions for the derivatives of pupil positions with respect to axial curvatures. Consider now the effect on the $\bar{O}T$ - coordinates under a change δp in p. Let $(\underline{S}, \underline{T})$ be the coordinates corresponding to the pupil position p and $(\underline{S}^*, \underline{T}^*)$ those corresponding to $p^* = p + \delta p$. Then

$$\underline{S} = \sigma\, Y_1 + \bar{\sigma}\, V_1, \quad \underline{I} = \tau\, \underline{Y}_1 + \bar{\tau}\, \underline{V}_1$$

where $\sigma = (1 - p/\ell_{01})^{-1}$, $\bar{\sigma} = -p\sigma$, $\tau = -1/\ell_{01}$, $\bar{\tau} = 1$

and $\underline{S}^* = \sigma^*\, \underline{Y}_1 + \sigma^*\, \underline{V}_1$, $\underline{T}^* = \tau^*\, \underline{Y}_1 + \bar{\tau}^*\, \underline{V}_1$

where $\sigma^* = 1/(1 - p^*/\ell_{01})$, $\bar{\sigma}^* = -p^*\sigma^*$, $\tau^* = \tau$, $\bar{\tau}^* = \bar{\tau}$.

Now
$$\underline{S} = \sigma^*\, \underline{Y}_1 + \bar{\sigma}^*\, \underline{V}_1 + (\sigma - \sigma^*)\, \underline{Y}_1 + (\bar{\sigma} - \bar{\sigma}^*)\, \underline{V}_1$$

$$= \underline{S}^* + (\sigma - \sigma^*)\, \underline{Y}_1 + (p^*\sigma^* - p\ )\, \underline{V}_1$$

Let $\chi = \ell_{01}(\sigma^* - \sigma)$, so $p^*\sigma^* - p\sigma = \chi$. Then we have

$$\underline{S} = \underline{S}^* - \chi/\ell_{01} \cdot \underline{Y}_1 + \chi_1 \underline{V}_1$$

i.e.
$$\underline{S} = \underline{S}^* + \chi \underline{T}^* \qquad (1.6.21)$$

Thus the change δp in p can be considered as a transformation of coordinates, if the coordinates are to have the same geometrical significance.

Now the aberration of a ray is a length, independent of any particular set of coordinates. Thus

$$\underline{\varepsilon}^{*\prime}_{k} = \underline{\varepsilon}^{\prime}_{k}$$

$$\text{so} \quad \sum_{n=1}^{\infty} \sum_{\mu=0}^{n} \sum_{\nu=0}^{\mu} (\overset{*}{G}{}^{\prime(n)}_{\mu\nu k} \underline{S}^{*} + \bar{\overset{*}{G}}{}^{\prime(n)}_{\nu\mu k} \underline{T}) . \overset{*}{\xi}{}^{n-\mu} . \overset{*}{\eta}{}^{\mu-\nu} . \overset{*}{\zeta}{}^{\nu} = \sum_{n=1}^{\infty} \sum_{\mu=0}^{n} \sum_{\nu=0}^{\mu} (G^{\prime(n)}_{\mu\nu k} \underline{S}$$

$$+ \bar{G}^{\prime(n)}_{\mu\nu k} \underline{T}) \xi^{n-\mu} . \eta^{\mu-\nu} \zeta .^{\nu} \qquad (1.6.22)$$

By substituting for $\underline{S}$, $\underline{T}$, ξ, η and ζ using (1.6.21), and then equating like terms the effect on the various coefficients of a shift in entrance pupil position can be found. The relevant primary, secondary and tertiary expressions are given in appendix 4. It should be noted that the effect of the change δp in p is required for both a- and b- primary and secondary coefficients but only for the a-tertiary coefficients.

§1.7 The Relation between Entrance Pupil Coefficients and Retardation Coefficients

This section is devoted towards giving an understanding of the transformation process involved in deriving relations between the retardation coefficients and the $\bar{O}$T coefficients. The methods given in OAC VI are exact for the coefficients of order less than ten in the retardation coordinates. However the transformation relations given involve the b-coefficients and hence it is not easy to relate the retardation coefficients π_i, σ_i, τ_i to the $\bar{O}$T coefficients p_j, s_j, t_j. Also the complexity of the algebra tends to obscure what is being done in the various stages of the process. After explaining the transformation, some approximations will be introduced; using these, exact relations between the π_i and p_j, and approximate relations between corresponding higher order coefficients will be given and their usefulness examined. These are used in chapter 3, where the predictions of different types of assessment functions are compared.

The $\bar{O}T$ coordinates are defined by

$$\underline{S} = \underline{Y}_1 - \underline{Y}_{pr} \quad , \quad \underline{T} = \underline{H}_1/\ell_{01}$$

where (Y_{pr}, Z_{pr}) are canonical coordinates of a ray from the object point $(\ell_{01}, -H_{y1}, 0)$ whose direction passes through the axial point of the paraxial entrance pupil, and (Y_1, Z_1) are the coordinates of any ray from the same object point. It is convenient to express the aberration expansion in terms of coordinates (ρ, θ, V) where

$$S_y = \rho\cos\theta, \qquad S_z = \rho\sin\theta, \qquad V = H_{y1}/\ell_{01}.$$

The aberration expansion is then written

$$\begin{aligned}\varepsilon_y = {} & p_1\rho^3\cos\theta + p_2(2+\cos2\theta)\rho^2V + (3p_3+p_4)\rho V^2\cos\theta + p_5V^3 \\ & + s_1\rho^5\cos\theta + (s_2+s_3\cos2\theta)\rho^4V + (s_4+s_6\cos^2\theta)\rho^3V^2\cos\theta + (s_7+s_8\cos2\theta)\rho^2V^3 \\ & \qquad + s_{10}\rho V^4\cos\theta + s_{12}V^5 \\ & + t_1\rho^7\cos\theta + (t_2+t_3\cos2\theta)\rho^6V + (t_4+t_6\cos^2\theta)\rho^5V^2\cos\theta + (t_7+t_8\cos2\theta \\ & \qquad + t_{10}\cos4\theta)\rho^4V^3 + (t_{11}+t_{12}\cos^2\theta)\rho^3V^4\cos\theta \\ & \qquad + (t_{15}+t_{16}\cos2\theta)\rho^2V^5 + t_{18}\rho V^6\cos\theta + t_{20}V^7 + O(9).\end{aligned} \tag{1.7.1}$$

$$\begin{aligned}\varepsilon_z = {} & p_1\rho^3\sin\theta + p_2\sin2\theta\rho^2V + (p_3+p_4)\rho V^2\sin\theta \\ & + s_1\rho^5\sin\theta + s_3\sin2\theta\rho^4V + (s_5+s_6\cos^2\theta)\rho^3V^2\sin\theta + s_9\sin2\theta\rho^2V^3 \\ & \qquad + s_{11}\rho V^4\sin\theta\end{aligned}$$

$$+ t_1\rho^7\sin\theta + t_3\sin 2\theta\rho^6 V + (t_5+t_6\cos^2\theta)\rho^5 V^2\sin\theta + (t_9\sin 2\theta + t_{10}\sin 4\theta)\rho^4 V^3$$
$$+ (t_{13}+t_{14}\cos^2\theta)\rho^3 V^4\sin\theta + t_{17}\sin 2\theta\rho^2 V^5 + t_{19}\rho V^6\sin\theta + O(9).$$

The retardation expansion is expressed in terms of coordinates (y,z,h) where (y,z) are coordinates on the actual wavefront and h is the paraxial image height.

The transformation process detailed in OAC VI and VII is as follows:-

a) The equivalent entrance pupil is mapped onto the paraxial exit pupil plane, the form of the mapping depending on the actual system being considered. Since rays passing through the system are aberrated the mapping is not linear but rather quasi-linear. Hence the transformation of coordinates involved is quasi-linear, and the effect on the coefficients is described by the transformation from $\bar{O}T$ to W coefficients detailed in 6 of OAC VI. The linear part of the transformation involves simply the paraxial magnification between pupil planes, m_{1e}, where

$$m_{1e} = (1-p/\ell_{01}).m_p = (1-p/\ell_{01})/(v'_{pk}.p + v'_{qk}),$$

p being the axial distance from the front polar tangent plane to the entrance pupil plane.

b) The transverse aberration must now be expressed in terms of coordinates on the actual wavefront, since the relationship

$$\underline{\varepsilon} = e\ \partial D/\partial_{\underline{y}}$$

is to be used. This involves a mapping of points on a plane onto a surface which is approximately a sphere, and this mapping is implicit in the relations of § 6 of OAC VII, which give the deformation coefficients in terms of the W coefficients. An approximation to the process would be a mapping of points on the exit pupil onto a sphere centred on the ideal image point such that corresponding points on the exit pupil plane and reference sphere are collinear with the ideal image point.

c) The final step is to take account of the relation between retardation and deformation functions. The form of the exact relations between these coefficients indicates that, in general, corresponding coefficients will be approximately equal though this is not necessarily so (e.g. systems corrected for large apertures tend to have spherical aberration coefficients of different orders which are of the same magnitude; the higher order retardation and deformation coefficients, under these conditions, will differ markedly. However the coefficient expansions are likely to have limited applicability for such systems).

In the use of aberration coefficients it is often satisfactory to have close approximations to the various coefficients which are easily calculated rather than precise values involving considerably more calculation. With this in mind the following approximate transformations were considered as substitutes for the exact methods explained above.

1) Project the equivalent entrance pupil onto the paraxial exit pupil according to paraxial optics. So if $S_y = \rho\cos\theta$, $S_z = \rho\sin\theta$, $T_y = V_1$, $T_z = 0$ are the $\bar{O}T$ coordinates of a ray then the approximate coordinates in the exit pupil plane are

$$S^*_y = e^{-1}m_{1e}\rho\cos\theta, \qquad S^*_z = e^{-1}m_{1e}\rho\sin\theta, \; T^*_y = h, \; T^*_z = 0$$

i.e. $\underline{S}^* = m_{1e}/e.\underline{S}$, $\qquad \underline{T}^* = m\ell_{01}\underline{T}$

2) Use the approximate mapping described in b) above, followed by the approximation, $\qquad R \approx (1-\tfrac{1}{2}e^{-2}h^2)D$

The resulting expressions relating (p,s,t) and (π,σ,τ) coefficients are exact for third order coefficients and are as follows

$$p_1 = 4e^4.m^{*-3}\pi_1, \qquad p_2 = e^3m^{*-1}\pi_2$$

$$p_3 = e^2m^*\pi_4, \quad p_4 = e^2m^*(2\pi_3-\pi_4), \quad p_5 = em^{*3}\pi_5 \tag{1.7.2}$$

In deriving corresponding expressions for the higher order coefficients it is necessary to make use of identities between the coefficients. In general these involve the b-coefficients of lower orders. However the identities between the W coefficients do not involve b-coefficients, and since under the approximation of (1) the W and $\bar{O}$T coefficients are related by multiplicative factors, the $\bar{O}$T coefficient identities can be replaced by the corresponding W coefficient identities. The error resulting from this approximation is that the constants multiplying lower order coefficients in a particular expression should in fact be expressions containing the constant and sums of the $\bar{O}$T b-coefficients. The fifth order approximate relations are

$$s_1 = e^6m^{*-5}.(\sigma_1-e^{-2}\pi_1),$$

$$s_2 = e^5m^{*-3}[3\sigma_2-e^{-2}(10\pi_1+2\pi_2)],$$

$$s_3 = e^5m^{*-3}.[2\sigma_2-e^{-2}(8\pi_1+\pi_2)],$$

$$s_4 = e^4m^{*-1}[4\sigma_3+2\sigma_4-e^{-2}.(2\pi_1+3\pi_2+\pi_4+\pi_5)],$$

$$s_5 = e^4m^{*-1}.[4\sigma_3-e^{-2}(2\pi_1+\pi_2+\pi_3)],$$

$$s_6 = e^4m^{*-1}[2\sigma_4-e^{-2}.(8\pi_1+4\pi_2)],$$

$$s_7 = e^3m^*.[2\sigma_5+\tfrac{3}{2}\sigma_7-e^{-2}.(5\pi_2+2\pi_3+\pi_4)],$$

$$s_8 = e^3m^*[\sigma_5+\tfrac{3}{2}\sigma_7-e^{-2}(\tfrac{7}{2}\pi_2+\pi_3+\pi_4)],$$

$$s_9 = e^3m^*.[\sigma_5-e^{-2}(\tfrac{3}{2}\pi_2+\pi_3)],$$

$$s_{10} = e^2m^{*3}[2(\sigma_6+\sigma_8)-e^{-2}.3(\pi_3+\pi_4)],$$

$$s_{11} = e^2m^{*3}[2\sigma_6-e^{-2}\pi_3],$$

$$s_{12} = em^{*5}[\sigma_9-\tfrac{1}{2}e^{-2}\pi_5] \tag{1.7.3}$$

The seventh order relations, neglecting third order coefficients in the identities, are,

$$t_1 = e^8 m^{*-7}(8\tau_1 - 15e^{-2}\sigma_1)$$

$$t_2 = e^7 m^{*-5}.4[\tau_2 - e^{-2}(\sigma_1+\sigma_2)]$$

$$t_3 = e^7 m^{*-5}[3\tau_2 - e^{-2}(21\sigma_1+4\sigma_2)]$$

$$t_4 = e^6 m^{*-3}[6\tau_3+2\tau_4 - e^{-2}(9\sigma_1-10\sigma_2+6\sigma_3+\tfrac{3}{2}\sigma_4)]$$

$$t_5 = e^6 m^{*-3}[6\tau_3 - e^{-2}(3\sigma_1+2\sigma_2+6\sigma_3)]$$

$$t_6 = e^6 m^{*-3}[4\tau_4 - e^{-2}(24\sigma_1+20\sigma_2+3\sigma_4)]$$

$$t_7 = e^5 m^{*-1}[3\tau_5+9/4.\tau_7 - e^{-2}(7\sigma_3+21/2.\sigma_2+35/4.\sigma_4+1/2.\sigma_5)]$$

$$t_8 = e^5 m^{*-1}[2\tau_5+5/2.\tau_7 - e^{-2}(\sigma_2+1/4.\sigma_3-\sigma_4-\sigma_5+3/2.\sigma_7)]$$

$$t_9 = e^5 m^{*-1}[2\tau_5+1/2.\tau_7 - e^{-2}(3\sigma_2+4\sigma_3-\sigma_4-\sigma_5)]$$

$$t_{10} = e^4 m^*[4\tau_6+2\tau_8 - e^{-2}(6\sigma_3+3\sigma_4+\sigma_5+\sigma_6+\sigma_8)]$$

$$t_{12} = e^4 m^*[2\tau_8+4\tau_{11} - e^{-2}(8\sigma_3+11\sigma_4+4\sigma_5+6\sigma_7)]$$

$$t_{13} = e^4 m^*[4\tau_6 - e^{-2}(2\sigma_3+\sigma_5+\sigma_6)]$$

$$t_{14} = e^4 m^*.[2\tau_8 - e^{-2}(8\sigma_3+5\sigma_4+4\sigma_5)]$$

$$t_{15} = e^3 m^{*3}.[2\tau_9+3/2.\tau_{12} - e^{-2}.(4\sigma_5+2\sigma_6+15/4.\sigma_7+2\sigma_8)]$$

$$t_{16} = e^3 m^{*3}.[\tau_9+3/2.\tau_{12} - e^{-2}.(7/2.\sigma_5+15/4.\sigma_7+\sigma_6+\sigma_8)]$$

$$t_{17} = e^3 m^{*3}.[\tau_9 - e^{-2}(3/2.\sigma_5+\sigma_6)]$$

$$t_{18} = e^2 m^{*5}[2(\tau_{10}+\tau_{13}) - 3e^{-2}(\sigma_6+\sigma_8)]$$

$$t_{19} = e^2 m^{*5}[2\tau_{10} - 1/2e^{-2}\sigma_6)]$$

$$t_{20} = em^{*7}(\tau_{14} - 1/2e^{-2}\sigma_9) \qquad (1.7.4)$$

The accuracy of the approximations in (1.7.3,4) has been numerically investigated for a number of systems. The results for three of these systems - the triplet Σ_1 used in L and OAC for numerical examples, a Topogon objective (as an example of a wide angle lens) and a Sonnar objective (for large aperture) are given in table 1.10 below. The results indicate that the simple approximations obtained by ignoring all third order coefficients in (1.7.3) and fifth order coefficients in (1.7.4) are more useful than their more complex counterparts. For these simple approximations the results indicate that the fifth order relations would be useful in general, except maybe for the coefficients s_{10}, s_{11} and s_{12} while the seventh order relations would only be useful for the coefficients t_1, t_2, t_3 and t_4.

	Σ_1			Topogon			Sonnar		
	True	Approx.	% Error	True	Approx.	% Error	True	Approx.	% Error
s_1	-90.924	-90,987	0.07	-8.720	-8.671	0.57	-6.662	-6.661	0.021
s_2	-69.244	-68.890	0.51	-2.000	-1.930	3.50	-0.3511	-0.3535	0.669
s_3	-46.071	-45.927	0.31	-1.327	-1.289	3.04	0.2394	0.2357	1.55
s_4	-20.631	-20.273	1.73	-3.371	-3.196	5.20	-0.6168	-0.5624	8.82
s_5	-8.760	-9.702	10.75	-2.251	-2.334	3.66	-1.976	-2.073	4.89
s_6	-13.188	-10.571	19.84	-1.326	-0.862	35.0	1.209	1.510	24.9
s_7	-1.736	1.896	9.23	-0.4021	-0.1871	53.47	-0.2903	-0.1215	58.1
s_8	1.023	1.071	4.74	-0.2366	-0.0335	85.8	-0.1238	-0.0357	71.2
s_9	0.749	0.825	10.18	-0.2233	-0.1536	31.2	-0.1684	-0.0859	49.0
s_{10}	-0.216	-0.175	18.88	-0.6194	-0.5031	18.8	-0.3841	-0.3771	1.82
s_{11}	-0.376	-0.468	24.44	-0.4127	-0.4333	5.00	-0.3020	-0.3158	4.59
s_{12}	-0.0645	-0.0469	27.39	-0.0169	0.0055	132	-0.1611	-0.1281	20.4
t_1	-4653.6	-4744.4	1.95	-93.79	-97.29	3.74	-58.75	-58.85	0.163
t_2	-2598.9	-2664.9	2.54	-66.58	-71.17	6.90	-4.478	-4.538	1.33
t_3	-1950.7	-1998.7	2.46	-50.08	-53.38	6.59	-3.746	-3.403	9.15
t_4	-647.6	-698.6	7.88	-48.63	-55.95	15.1	-14.05	-15.75	12.0
t_5	-324.0	-266.1	17.87	-23.38	-20.97	10.3	-20.10	-16.12	19.8
t_6	-506.1	-864.9	70.89	-34.73	-69.97	101	23.42	0.7556	96.8

t_7	-46.00	-221.4	380.1	-1.296	-11.49	787	2.435	-6.370	362
t_8	-21.94	-226.0	930	0.4232	-8.862	1000	2.315	-3.139	236
t_9	0.5841	-69.11	1000	-1.265	-6.452	410	0.6091	-5.354	979
t_{10}	2.326	-19.62	944	0.3538	0.3012	185	0.1995	0.2769	38.8
t_{11}	-25.04	-47.04	87.8	-3.425	-6.878	107	0.1938	0.4027	108
t_{12}	4.047	-74.79	1000	4.786	-5.409	213	1.031	2.782	170
t_{13}	-4.550	-3.073	32.46	-2.047	-1.562	23.7	-1.311	-0.0359	97.3
t_{14}	1.807	-43.96	1000	1.857	-5.316	386	2.269	0.4386	80.7
t_{15}	3.535	1.485	58.0	0.4718	-1.179	350	-0.6487	-0.4577	29.4
t_{16}	2.379	2.159	9.25	0.1787	-0.7574	524	-0.1028	0.0133	113
t_{17}	0.312	0.704	125	0.2007	0.1763	12.2	-0.1003	0.0973	198
t_{18}	1.771	1.991	12.34	0.4500	0.0107	97.6	0.5677	0.172	69.8
t_{19}	0.091	0.477	426	0.1713	0.3322	93.9	0.2231	0.2573	15.3
t_{20}	0.0719	0.0867	20.6	0.0377	-0.0307	187	0.1411	0.0792	86.4

Table 1.10

The $\bar{O}T$ coefficients have been introduced since these are to be used in chapter 2 to derive expressions for a number of assessment functions based on the spot diagram. However it should be noted that an aberration expansion similar to (1.7.1) could be given using coordinates defined on the gaussian reference sphere. Sands has introduced a method which allows the calculation of the transverse aberration coefficients using such coordinates - called W_o coordinates - but these coordinates have not been used in this work. Using the W_o coordinates and an appropriate aberration expansion the relations between wavefront retardation coefficients and the transverse aberration coefficients are easily found using (1.4.3) and a transformation from coordinates (y_o, z_o) on the reference sphere and (y,z) on the actual wavefront.

It is a common procedure to determine the wave aberration by tracing a set of rays through specified points in the entrance pupil and using the wavefront aberration values thus found to fit a polynomial of specified form to these values. In this approach the wavefront aberration is expressed in terms of entrance pupil coordinates. The coefficients in this type of polynomial are approximate coefficients of the aberration expansion, the accuracy of the approximation depending on how well the true expansion, when truncated to the form of the chosen polynomial represents the actual aberration function. The above results indicate that the approximate coefficients found by ray-tracing and the analytic coefficients cannot be satisfactorily compared in general, except through use of the exact relationships, involving b-coefficients, between entrance pupil and exit pupil coefficients.

References

1.1) M. Herzberger, Modern Geometrical Optics (Interscience Publishers, New York, 1958), p.383.

1.2) W.B. King, Appl. Opt., 7, 489 (1968).

1.3) H.A. Buchdahl, Optical Aberration Coefficients (Dover Publications, Inc., New York, 1968).

1.4) P.W. Ford, Thesis, University of Tasmania (1962); P.J. Sands, Thesis, Australian National University (1967); H.A. Buchdahl, J.O.S.A., 49, 1113 (1959).

1.5) H.A. Buchdahl, J.O.S.A., 50, 539 (1960).

1.6) H.A. Buchdahl, An Introduction to Hamiltonian Optics (Cambridge University Press, 1970), p.36.

1.7) P.J. Sands, Appl. Opt., 9, 828 (1970).

1.8) H.A. Buchdahl, J.O.S.A., 50, 534 (1960).

1.9) M. Born, and E. Wolf, Principles of Optics (Pergamon Press, Inc., New York, 1968), p.205.

1.10) H.A. Buchdahl, ref. 1.6, p.106.

1.11) J.L. Rayces, Optica Acta, 11, 85 (1964).

1.12) H.A. Buchdahl, ref. 1.4.

1.13) P.J. Sands, ref. 1.4.

CHAPTER TWO

IMAGE QUALITY ASSESSMENT FUNCTIONS

§2.1 Introduction

To systematically carry out the design of an optical system it is necessary to have some quantity which assigns a numerical value to a particular state of correction, this value being greater than that for any other state which is less satisfactory for the use to which the system is to be put. Such a quantity, which orders the states of correction of optical systems according to their supposed usefulness for a particular task is called an assessment function. When an optical system is assessed for image quality it is done using information obtained from a knowledge of the form of the image at each of a discrete set of object points, whether this information be that given by geometrical optics or scalar diffraction theory. In the following work the assessment functions considered are only applied to the aerial image formed by the optical system. A brief survey is given of significant assessment quantities and of work on their relationships to one another. This is followed by the introduction of some new quantities, based on the work of Dutton and defined in terms of the optical transfer function (o.t.f.). It is shown that analogous quantities can be defined for the spot diagram (s.d.) and relationships between the two types of quantities are derived. Some additional assessment functions are introduced and methods are given for the computation of these assessment functions, using aberration coefficients.

§2.2 Historical Survey

Longitudinal and transverse aberrations of rays were well-established representations of the state of correction of an optical system prior to the work of Conrady.[1] The coefficients of variables in the power series expansion for the aberration of a ray, the aberration coefficients, were much used followi the work of Seidel.[2] Herzberger developed the use of the spot diagram representation of image quality and introduced the diapoint concept to aid in the analysis of aberration types. Cruickshank and Hills[3] also used spot diagram analysis, with higher order aberration coefficients, to systematize design work. Quantitative use of the spot diagram was introduced by Hawkins and Linfoot[4] through the use of the average of the squared aberration across the entrance pupil. This quantity has the disadvantage that rays in the outer part of the image contribute very substantially to the value of the function while quite probably having negligible effect on the image actually detected apart from a reduction of contrast. To obviate this difficulty, Lucy[5] proposed the quantity

$$L = A^{-1} \iint (|\varepsilon| + K)^{-1} \, dA$$

where K is a positive constant, such as the radius of the Airy disk for the aperture A.

Concern for diffraction effects at the stop of the system led naturally to consideration of the geometrical wavefront aberration and quantities based on this. Marechal[6] showed that, for well-corrected systems, the r.m.s. wavefront aberration was directly proportional to the Strehl ratio. A number of workers[7,8] tackled the problem of determining the effects of particular aberration types on the scalar diffraction patterns of point objects (the point spread function or p.s.f.). However the difficulty of computation and of determining the effects of aberration types on the p.s.f. resulted in there being no significant use of this quantity during optical design.

The introduction of the o.t.f. by Duffieux[9] and Schade[10] made consideration of diffraction effects a far more feasible proposition than was the use of the point spread function. The use of the o.t.f., and quantities based on it, was significantly advanced by Hopkins' school at Imperial College. With the introduction of the geometrical transfer function as an approximation to the o.t.f.[11,12] and the development of further approximations which could be rapidly computed the use of spatial frequency concepts in design and assessment was assured. King[13] considered the computation of quantities such as the variance of the aberration difference function which arises in the transfer function integral, while Hopkins[14] used an o.t.f. approximation to derive tolerance criteria for aberration types which were frequency dependent.

Considerations of an optical system as a device which transmits information led Linfoot[15] to develop a number of quantities defined in terms of the o.t.f. which measure certain properties of the image such as sharpness and fidelity. Recently Lanzl[16] has introduced a quantity called the self-entropy of an image which, though rather tedious to calculate, may be rapidly measured experimentally.

Considerable attention has been paid to the calculation of the fraction of total energy transmitted from a given object point which lies within a circle of specified radius in the image plane. For photographic objectives which are not approaching the diffraction limit, Stavroudis[17] has shown that the circle within which 30% of the energy lies gives a good indication of photographic resolution. Expressions for the encircled energy in the case of apodization have been given by De.[18]

Dutton[19] has proposed a set of spatial frequency functions, determined from measured o.t.f.'s, which give a useful measure of some image defects. He uses the geometric mean of the modulation transfer function in radial and tangential sections to measure image spread, the fractional differences in m.t.f. for radial and tangential sections to measure "astigmatism" and the imaginary part of the o.t.f. to measure spread function assymetry. It should be noted that the defects measured by Dutton are not independent since asymmetry affects the radial symmetry of the image (departures from which have been classified as "astigmatism"), and both asymmetry and "astigmatism" imply image spread. However attention to the image in the context of these three separate classifications appears to be implicit in most design work.

With the proliferation of image assessment quantities there arose the need for studies of the relations between them. Wolf[20] derived the well-known differential relation between ray and wave aberrations - though it should be noted that Hamilton had established such a relation between the mixed characteristic and ray aberrations at a much earlier date. Hopkins used this result in deriving the g.t.f., which related the transverse aberrations, and the spot diagram, to the scalar diffraction pattern. Miyamoto extended this analysis in a series of papers[21] in which he showed that, for aberration terms which gave a maximum aberration within the pupil of greater than about 2λ, there was no significant difference between the g.t.f. and o.t.f. This analysis was carried out for third order aberrations and defocussing. Hopkins established empirically that, provided the modulus of the transfer function of a particular system was greater than about 0.7 of the corresponding response for an aberration free system then the variance of the wave aberration difference function was a useful criterion of image quality. No satisfactory relationship has been found between the o.t.f. and geometrical quantities for systems which neither approach the diffraction limit, nor are sufficiently well-corrected for the g.t.f. to be suitable,nor satisfy Hopkins' cirterion, mentioned above, over a significant frequency range. Kapany and Burke[22] have also established relations between various diffraction based assessment quantities.

Little attention has been paid to the differences in assessment of a particular system using different criteria. However De Velis[23] has made an interesting study of this for the case of defocussing. Using minimization of the radius of gyration of spot diagrams to establish an optimum plane for geometrical optics and minimization of wavefront variance for physical optics, it was found that the two planes so determined defined bounds within which the optimum plane according to various other criteria based on the o.t.f. lay. Apart from this, there is work by the Japan Optical Engineering Research Association[24], carried out on fabricated systems, which has shown that different criteria may lead to different orderings of a given set of systems.

From the foregoing it can be seen that, although the rapid computation of diffraction-based image quality criteria has been thoroughly examined, the comparative effects of using different criteria have not been determined sufficiently. Since, as mentioned earlier, a considerable amount of design is done by workers with only limited facilities, it is clear that some better idea of the differences in assessment using different criteria is needed. The writer, therefore, has devoted the remainder of this chapter and the following chapter to an investigation of these differences. The differences we are concerned with are those between the assessments of image quality provided by a certain set of functions on a given optical system. However by choosing to investigate a range of systems of widely varying image quality the dependence of any such differences in assessment on the image quality can also be studied.

§2.3 New Assessment Functions

Let $T(R, \theta)$ denote the normalized frequency response of an optical system at a line frequency R, for lines at an orientation θ to the tangential section. Let $r = 2\tilde{e}R$, where $\tilde{e}$ is the distance from the axial point of the entrance pupil to the image point, and let $s = \lambda\, r \sin\theta$, $t = \lambda\, r \cos\theta$, λ being the wavelength of light. Then

$$T(R,\theta) = A^{-1}_{(o,o)} \cdot \iint_{A(s,t)} \{ \exp [\, jk\, W(y+s,\ z+t) - W(y-s,\ z-t)]\}\, dydz \tag{2.3.1}$$

where $A(s,t)$ denotes the region common to replicas of the pupil with centres at $(\pm s, \pm t)$ and $W(y,z)$ is the wavefront aberration function. The three functions are then defined by

$$S_1(R) = \mathrm{Re}\ \{T(R, \tfrac{\pi}{2})\} \cdot \ \mathrm{Re}\ \{T(R,0)\}$$

$$S_2(R) = \mathrm{Re}\ \{T(R, \tfrac{\pi}{2}) - T(R,0)\}$$

$$S_3(R) = \mathrm{Im}\ \{T(R, \tfrac{\pi}{2}) - T(R,0)\}$$

The functions used by Dutton are, apart from S_3, analogous quantities defined using the modulus and phase rather than real and imaginary form of the o.t.f.. $S_1(R)$ provides a measure of how sharp the image is, while $S_2(R)$ measures the difference between the contrast in objects of the same frequency lying at orientations of $\frac{\pi}{2}$ to one another. The presence of imaginary terms in the o.t.f. indicates asymmetry and so $S_3(R)$ provides a measure of asymmetry.

Let $\underline{\varepsilon}(\rho,\theta)$ denote the aberration of a ray passing through the point (ρ,θ) in the entrance pupil from a given object point. The aberration is referred to a Cartesian coordinate system in the ideal image plane with origin at the gaussian image point. Let $(\varepsilon_y, \varepsilon_z)$ be the components of $\underline{\varepsilon}$.

The three spot diagram quantities are then defined as follows:

$$P_1 = A^{-1} \iint_A (\varepsilon_y^2 + \varepsilon_z^2)\ da$$

$$P_2 = A^{-1} \iint_A (\varepsilon_y^2 - \varepsilon_z^2)\ da \qquad (2.3.3)$$

$$P_3 = A^{-1} \iint_A \varepsilon_y\ da$$

By referring the aberration to an origin at $(-P_3, 0)$ more suitable values for P_1 and P_2 may be found.

§2.4 Relations Between the Two Types of Quantities

Geometrical optics can be considered as a limiting case of physical optics as the wavelength, λ, tends to zero. Thus a geometrical transfer function is defined by considering (2.3.1) in the limit as $\lambda \to 0$.

$$Tg(R,\theta) = A^{-1} \iint_A \exp\{j\ \overset{4}{\cancel{2}}\pi r(\frac{\partial W}{\partial y} \sin\theta + \frac{\partial W}{\partial z} \cos\theta)\}\ dydz \qquad (2.4.1)$$

Now $\underline{\varepsilon} = e \dfrac{\partial D}{\partial \underline{y}} (y,\ z,\ h)$ (1.4.3)

where $\underline{\varepsilon}$ is the transverse ray aberration, e is the axial distance from the paraxial exit pupil to the gaussian image plane, and D is the deformation function. Also

$$W = (1 - \tfrac{1}{2} e^{-2}h^2 + \frac{3}{8} e^{-4}h^4)\ D + \tfrac{1}{2} e^{-1}D^2 + 0(10) \qquad (1.4.5)$$

hence $\dfrac{\partial W}{\partial \underline{y}} = \phi \dfrac{\partial D}{\partial \underline{y}} + 0(9) = \phi e^{-1}\underline{\varepsilon} + 0(9)$ (2.4.2)

where $\phi = 1 - \tfrac{1}{2} e^{-2}h^2 + \frac{3}{8} e^{-4}h^4 + e^{-1}\ D$ (2.4.3)

Substituting for $\dfrac{\partial W}{\partial \underline{y}}$ in (2.4.1) using (2.4.2) gives

$$Tg(R,\theta) = A^{-1}\iint \exp\{j\ 2\pi r\phi e^{-1}(\varepsilon_y \sin\theta + \varepsilon_z \cos\theta + 0(9))\}\ dydz \qquad (2.4.4)$$

This is an exact relation for the frequency response according to geometrical optics.

Suppose that the aberration associated with any ray is adequately represented by terms of order less than nine in the expression for $\underline{\varepsilon}$. Let $\alpha = \overset{4}{\cancel{2}}\pi e^{-1}$ and $\chi(\theta) = \varepsilon_y \sin\theta + \varepsilon_z \cos\theta$. Then

$$Tg(R,\theta) = A^{-1} \iint \exp\{j\ \alpha r\phi\chi\}\ dydz \qquad (2.4.5)$$

Expanding the exponential in (3.4.5) as a power series gives

$$Tg(R,\theta) = A^{-1}\iint[1-\alpha^2r^2\phi^2\chi^2(\theta) + \alpha^4r^4\phi^4\chi^4(\theta) + 0(18)]\,dydz$$
$$+ j\,A^{-1}\iint\{\alpha r\phi\chi(\theta)\,[1-\alpha^2r^2\chi^2(\theta)] + 0(12)\}\,dydz \qquad (2.4.6)$$

so that

$$S_1g(R) = 1-\alpha^2r^2A^{-1}\iint\phi^2(\varepsilon_y{}^2 + \varepsilon_z{}^2)\,dydz + \alpha^4r^4A^{-1}\iint\phi^4(\varepsilon_y{}^4 + \varepsilon_z{}^4)\,dydz$$
$$+ \alpha^4r^4A^{-2}\iint\phi^2\varepsilon_y{}^2\,dydz \iint\phi^2\varepsilon_z{}^2\,dydz + 0(18) \qquad (2.4.7)$$

From (2.4.7) it can be seen that the series expansion for $S_1g(R)$ is rapidly convergent either for small r or small aberrations. However $r = 2\,R\,e = 2\lambda^{-1}.R/R_{max}$ so the criterion for rapid convergence is $\alpha^2\,\overline{\varepsilon^2}\,r^2 << 1$. This gives

$$(\overline{\varepsilon^2})^{\frac{1}{2}} < \frac{e}{8\cancel{4}\pi(R/R_{max})}\,\lambda \qquad (2.4.8)$$

Now $\phi = 1 - 0.5\,e^{-2}\,h^2 + 0.375\,e^{-4}h^4 + e^{-1}\,D$, so at a given field angle ϕ is approximately constant. Let

$$\phi = \psi + e^{-1}\,D = \psi + 0(4)$$

where ψ is constant at a given angle. Then

$$\phi\varepsilon = \psi(\varepsilon^{(3)} + \varepsilon^{(5)} + \varepsilon^{(7)}) + e^{-1}\,D^{(4)}\,\varepsilon^{(3)} + 0(9) \qquad (2.4.9)$$

If only third and fifth order aberrations are significant then

$$1 - S_1g(R) = \alpha^2\psi^2r^2A^{-1}\iint(\varepsilon^{(3)} + \varepsilon^{(5)})^2\,dydz + 0(12) \qquad (2.4.10)$$

that is $1 - S_1g(R) \propto P_1$. Equation (2.4.10) establishes the link between S_1 and P_1.

Often a geometrical transfer function is defined by

$$\overline{T}(R) = \int_0^{2\pi} Tg(R,\theta)\,d\theta$$

Using the same approximations as above this reduces to

$$\overline{T}(R) = \overset{4}{\cancel{2}}\pi\,S_1g(R)$$

so that minimising the radius of gyration of the spot diagram leads to maximising of the g.t.f.

Suppose, however, that $\varepsilon^{(7)} \simeq \varepsilon^{(3)} + \varepsilon^{(5)}$ and $D^{(4)}\,\varepsilon^{(3)} \simeq \varepsilon^{(7)}$ then (2.4.10) is not valid. This situation is quite feasible in practice, for suppose a design is carried out by balancing third and fifth order aberrations against one another – a reliable method when higher order terms are going to be fairly small. Then it may well be that $D^{(4)}\,\varepsilon^{(3)} \simeq \varepsilon^{(7)}$. This possibility arises from consideration of the effect of the term ϕ in

$$\underline{\varepsilon} = \phi\, e^{-1} \frac{\partial W}{\partial y} + O(9)$$

indicating that neglect of the variation in ϕ with ray coordinates may introduce significant errors under some circumstances.

By considerations similar to those leading to (2.4.10) it may be shown that

$$S_2 g(R) \rightarrow \alpha^2 r^2 A^{-1} \iint_A \phi^2 (\varepsilon_y{}^2 - \varepsilon_z{}^2)\, dydz + O(\varepsilon^4 r^4) \qquad (2.4.11)$$

$$S_3 g(R) \rightarrow \alpha r A^{-1} \iint_A \phi \varepsilon_y\, dydz + O(\varepsilon^2 r^2) \qquad (2.4.12)$$

Hence the quantities S_1, S_2, S_3 and P_1, P_2, P_3 may be considered analogous quantities for the assessment of correction states in physical optics and geometrical optics respectively.

§2.5 Calculation of the Assessment Functions

The calculation of the above quantities requires a knowledge of the form of the entrance or exit pupil and the transverse or ray aberration associated with any ray passing through the pupil. The aberration may be found by tracing a sufficiently large number of rays through the system and using a summation method to determine the various functions, or by choosing some form of aberration polynomial to describe the aberration and determining the coefficients by least-squares fitting, or else by using analytic expressions for the coefficients of a power series approximation to the aberration function. The last-mentioned method is used here, thus allowing us to calculate 28 coefficients which describe the third, fifth and seventh-order transverse ray aberration for a given system and specified object plane, and likewise for the fourth, sixth and eighth order wavefront aberration. King[25] has described methods for choosing rays to be used in determining the aberration by the second method, the resulting coefficients being valid for a specified object point. For the first method some idea of what a "sufficiently large number" of rays required for a stable representation of the quantities P_1, P_2 and P_3 is given by the accompanying graphs (fig. 2.1) determined for a simple triplet objective. It is clear that increasing the angle leads to an increase in the number of rays required,

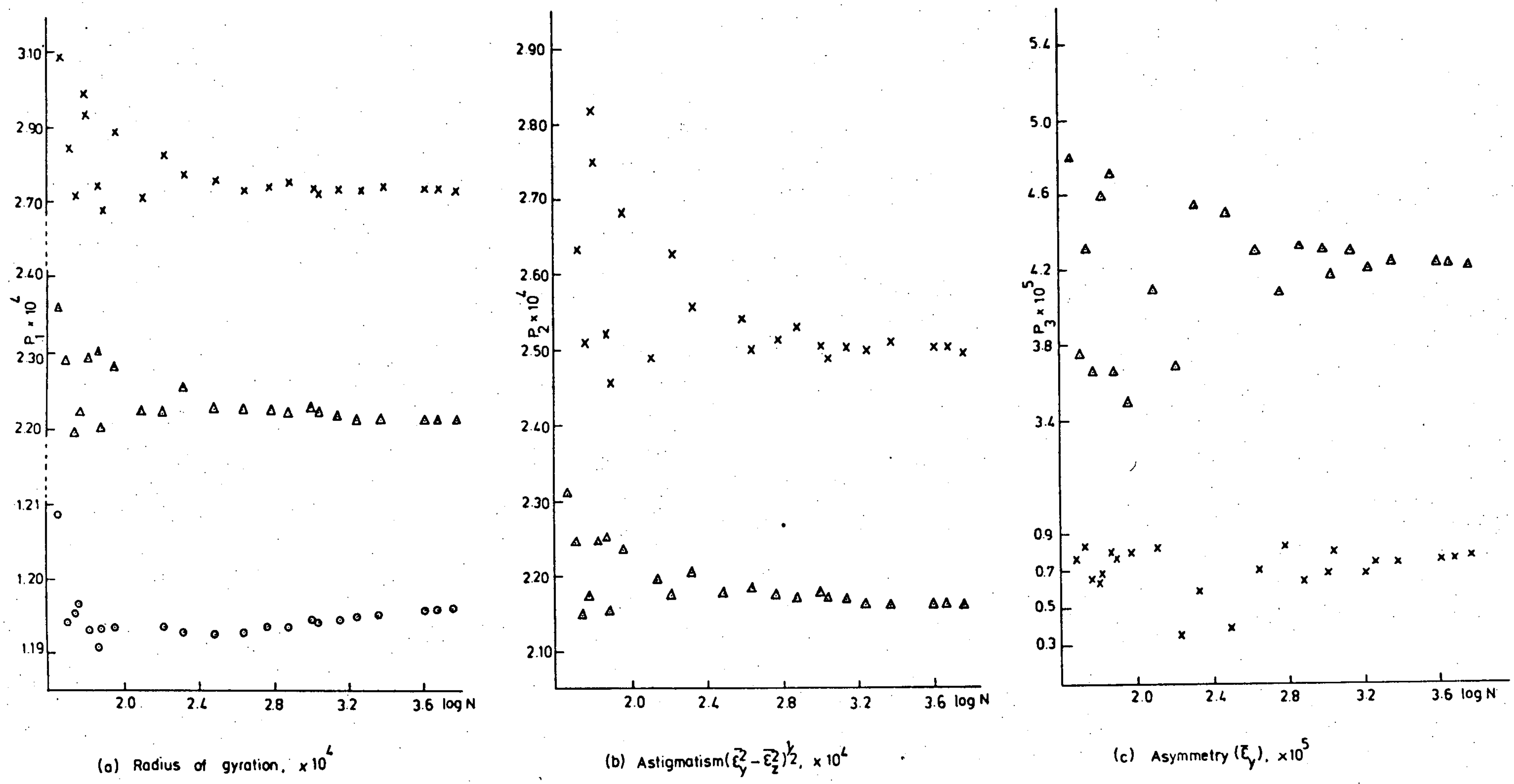


fig. 2.1 Stability of spot diagram functions, $P_j \sim \log$ (number of rays). $\circ \rightarrow$ axial, $\Delta \rightarrow 7^\circ$, $\times \rightarrow 10.5^\circ$.

and that the assymetry calculation is far more unstable than the larger-valued P_1 and P_2 quantities. This instability raises considerable problems for the calculation of numerical derivatives of these quantities, which do not arise when using the coefficients method.

§2.5(a) Determination of the Vignetted Pupil:

Realizing that paraxial optics fails to predict adequately the vignetted pupil most designers resort to some iterative raytrace to define the pupil periphery. Following earlier work on determination of pupil curves and peripheries by inversion of aberration expansions,[26] Sands[27] has proposed a method based on the paraxial and third-order coefficients which allows one to predict whether a given ray will be vignetted at any of the surfaces of the system. Three methods of pupil determination developed from Sand's prediction of vignetting have been used to calculate pupil peripheries for a number of photographic objectives, and the results have been compared with those obtained using an iterative finite raytrace method. It is found that an iterative "predicted raytrace" method is quite accurate, reliable and fast and requires very little computer storage space.

Following Sands we can write

$$\chi = y^2 + z^2 = [\xi + x^2(c\zeta + \eta)]/(1 - c\eta) \qquad (2.5.1)$$

where (x,y,z) are coordinates of a point on a spherical surface,

$$\xi = Y^2 + Z^2 , \qquad \eta = Y V + ZW, \qquad \zeta = V^2 + W^2$$

Now (Y,Z,V,W) can be expressed as a power series in the initial coordinates (Sy, Sz, Ty, Tz) of the ray. Sands gives two approximations for χ in terms of these initial coordinates - a modified paraxial approximation

$$\chi = ({y_a}^2\overline{\xi} + 2y_a y_b\overline{\eta} + {y_b}^2\overline{\zeta})/\{1-c[y_a v_a\overline{\xi} + (y_a v_b + y_b v_a)\overline{\eta} + y_b v_b\overline{\zeta}]\} \qquad (2.5.2)$$

where $\overline{\xi} = Sy^2 + Sz^2$, $\overline{\eta} = SyTy + SzTz$, $\overline{\zeta} = Ty^2 + Tz^2$

and an expression using the third order aberration coefficients for the surfaces of the system preceeding the one under discussion,

$$\chi = p_1\overline{\xi} + p_2\overline{\eta} + p_3\overline{\zeta} + s_1\overline{\xi}^2 + s_2\overline{\xi}\overline{\eta} + s_3\overline{\xi}\overline{\zeta} + s_4\overline{\eta}^2 + s_5\overline{\eta}\overline{\zeta} + s_6\overline{\zeta}^2, \qquad (2.5.3)$$

the p_i and s_j being given by Sands.

Suppose the entrance pupil is required for a system having an object plane at infinity. Then $\overline{OT}$ coordinates are used, so

$$\bar{\xi} = \rho^2, \qquad \bar{\eta} = V_1 \rho \cos\theta , \qquad \bar{\zeta} = V_1{}^2$$

where ρ is the radius of the entrance pupil and V_1 is the direction tangent of the ray from the object point aimed at the centre of the paraxial exit pupil. To determine the periphery, equations (2.5.2) or (2.5.3) are solved for ρ using a range of values of θ from 0° to 180° and with $x = \rho_d{}^2$ where ρ_d is the radius of the vignetting aperture. From (2.5.2)

$$\rho = [-\beta\, V_1 \cos\theta \mp \sqrt{\beta^2\, V_1{}^2 \cos^2\theta - 4\alpha\gamma}\,]/2\alpha \qquad (2.5.4)$$

where $\alpha = y_a\,(c\, v_a\, \chi + y_a)$, $\beta = [c(y_a\, v_b + y_b\, v_a)\chi + 2y_a y_b]$,

$$\gamma = y_b\; V_1{}^2(c\; v_b\; \chi + y_b) - \chi$$

By considering the case when $\theta = 0°$ and the stop is at the first surface it can be seen that the positive square root should be taken. This value of ρ together with V_1 and θ define a marginal modified-paraxial ray for the surface under consideration.

If it is assumed that the variables of a finite ray are adequately represented by the paraxial and primary coefficients and that paraxial optics gives a reasonable first approximation to the path of a ray then a solution of (3.7.3) in the region of the corresponding colution to (2.5.2) will give the initial coordinates of a marginal finite ray. Thus (2.5.3) is solved by some iterative method using the solution to (2.5.2) as an initial estimate.

Vignetting at all surfaces may be taken account of by using a fixed set of values of θ and finding ρ for this θ for all surfaces. The minimum value of ρ is then the coordinate of a peripheral point.

The iterative predicted ray trace is precisely the same as an iterative finite raytrace method except that rather than trace an actual ray and check on vignetting at each vignetting aperture the modified paraxial or first order method of predicting vignetting is used. Let (r,θ) denote the pupil coordinates of some ray, and consider the iteration process for a ray in this section. Suppose the vignetting aperture has a radius of ρ_j and the ray (r_i,θ) meets the surface at a radius ξ_j. If $\xi_j \leqslant \rho_j$ then the ray is transmitted otherwise it is vignetted and a new value for r, r_{i+1} must be chosen. This can be done either by decreasing r in fixed steps $r_{i+1} := r_i - \alpha\, r_o$, where α is a fraction determined by accuracy considerations, or else by using the difference between ξ_j and ρ_j

$$r_{i+1} := r_i - \beta\, \varepsilon_j \quad \text{where} \quad \varepsilon_j = \xi_j - \rho_j$$

and β is a factor chosen by consideration of convergence rates. The first method is obviously reliable and slow, while the second method may appear unreliable. However tests run using this method have been reliable and tolerances of 0.1% on the convergence have been used with need for no more than 6 iterative steps per ray when $\beta = 0.75$. Use of (2.5.4) is computationally sound, but (2.5.3), though more accurate than (2.5.4) in general, becomes unreliable when two solutions of the ~~cubic~~ quartic equation lie near one another. Hence its use is not recommended.

All the above methods were programmed for the Hydro-University Elliott 503 computer and a number of different photographic objectives were used as test data. Vignetting was assumed to occur at any surface aperture and a number of peripheral points calculated. The accuracy of the methods can be seen from the graphs which plot lower and upper extremities of the pupil against field angle. The two modified paraxial methods take almost the same time, this being no more than 1/10th the time taken by the finite ray trace method.

The exit pupil may be determined from the vignetted entrance pupil by paraxial projection or by using the third-order coefficients to find the intersection of the peripheral rays with the exit pupil surface. However neither of these methods has the accuracy of the entrance pupil determination, and the method recommended is to use a finite ray trace adjustment following entrance pupil determination using one of the modified paraxial methods above.

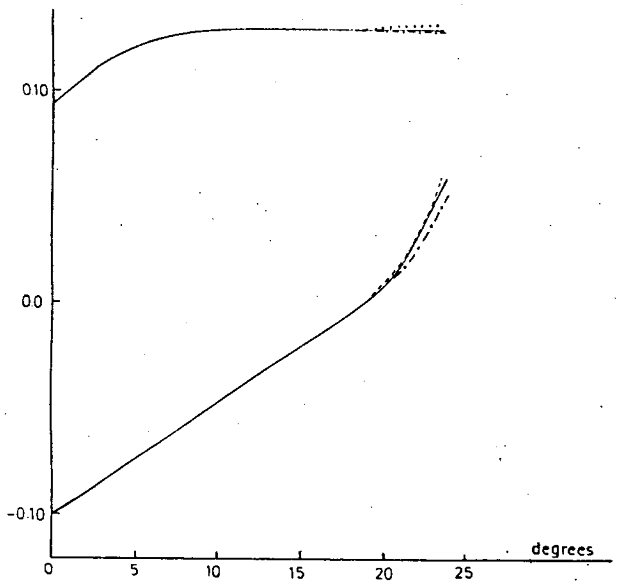


fig. 2.2 Upper and lower pupil extremities as a function of field angle. Petzval objective lens system.

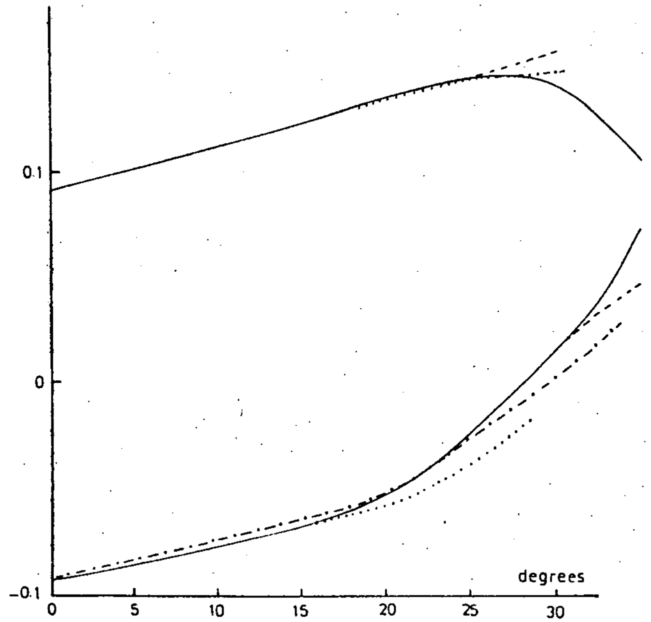


fig. 2.3 Triplet lens system.

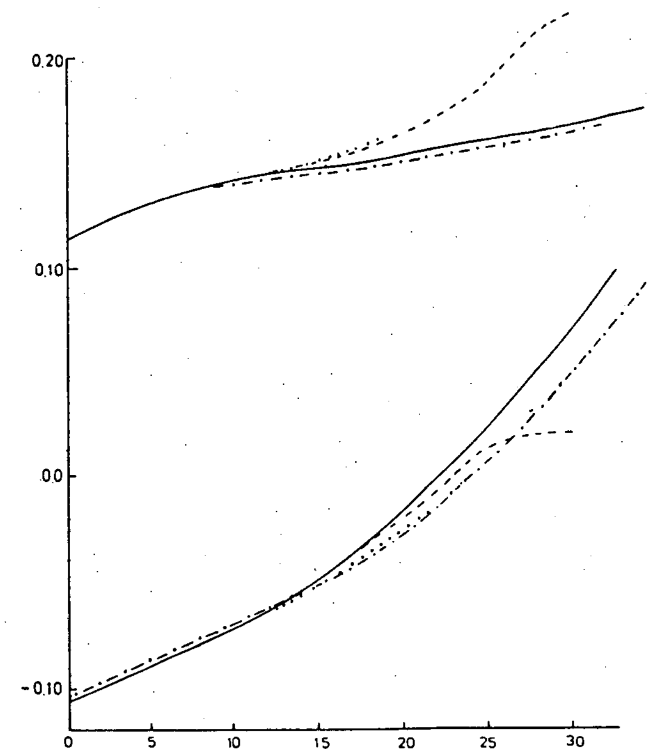


fig. 2.4 Topogon objective.

——— ray trace; ...... iterative paraxial pseudo-trace; —.—.— modified paraxial inversion; ------ third order inversion.

§2.5(b) Calculation of Spot Diagram Quantities:

Let $(\varepsilon_y, \varepsilon_z)$ be the components of the aberration of a ray from the point $(\ell o1, -H, 0)$ and passing through the point (ρ,θ) in the entrance pupil. Let $V = H/\ell o1$. Using Buchdahl's notation write

$$\varepsilon_y = \sigma_1\rho^3\cos\theta + \sigma_2(2+\cos 2\theta)\rho^2V + (3\sigma_3+\sigma_4)\rho V^2\cos\theta + \sigma_5V^3$$
$$+\mu_1\rho^5\cos\theta+(\mu_2+\mu_3\cos 2\theta)\rho^4V+(\mu_4+\mu_6\cos^2\theta)\rho^3V^2\cos\theta+(\mu_7+\mu_8\cos 2\theta)\rho^2V^3$$
$$+\mu_{10}\rho V^4\cos\theta+\mu_{12}V^5$$
$$+\tau_1\rho^7\cos\theta+(\tau_2+\tau_3\cos2\theta)\rho^6V+(\tau_4+\tau_6\cos^2\theta)\rho^5V^2\cos\theta+(\tau_7+\tau_8\cos2\theta+\tau_{10}\cos4\theta)$$
$$\rho^4V^3+(\tau_{11}+\tau_{12}\cos^2\theta)\rho^3V^4\cos\theta+(\tau_{15}+\tau_{16}\cos2\theta)\rho^2V^5+\tau_{18}\cos\theta\rho V^6+\tau_{20}V^7,$$

$$\varepsilon_z = \sigma_1\rho^3\sin\theta+\sigma_2\rho^2V\sin2\theta+(\sigma_3+\sigma_4)\rho V^2\sin \tag{2.5.5a}$$
$$+\mu_1\rho^5\sin\theta+\mu^3\rho^4V\sin2\theta+(\mu_5+\mu_6\cos^2\theta)\rho^3V^2\sin\theta+\mu_9\rho^2V^3\sin2\theta+\mu_{11}\rho V^4\sin\theta$$
$$+\tau_1\rho^7\sin\theta+\tau_3\rho^6V\sin2\theta+(\tau_5+\tau_6\cos^2\theta)\rho^5V^2\sin\theta+(\tau_9\sin2\theta+\tau_{10}\sin4\theta)\rho^4V^3$$
$$+(\tau_{12}+\tau_{14}\cos^2\theta)\rho^3V^4\sin\theta+\tau_{17}\rho^2V^5\sin2\theta+\tau_{19}\rho V^6\sin\theta \tag{2.5.5b}$$

In order to simplify the manipulation and integration of ε_y and ε_z it is convenient to write these in the form

$$\varepsilon_y = \overline{A} + \overline{B}\cos\theta + \overline{C}\cos^2\theta + \overline{D}\cos^3\theta + \overline{E}\cos^4\theta \tag{2.5.6}$$
$$\varepsilon_z = A + B\sin\theta + C\cos\theta\sin\theta + D\cos^2\theta\sin\theta + E\cos^3\theta\sin\theta$$

Now write $X = \sum_{i=0}^{7} x_i\rho^i$ for $X = A,B, \ldots ,\overline{A},\ldots\overline{E}$. Then the non-zero quantities are as follows,

$$a_0 = \sigma_5V^3 + \mu_{12}V^5 + \tau_{20}V^7$$
$$a_2 = \sigma_2V + (\mu_7-\mu_8)V^3 + (\tau_{15}-\tau_{16})V^5$$
$$a_4 = (\mu_2-\mu_3)V + (\tau_7-\tau_8+\tau_{10})V^3$$
$$a_6 = (\tau_2-\tau_3)V$$

$$b_1 = (3\sigma_3+\sigma_4)V^2 + \mu_{10}V^4 + \tau_{18}V^6$$
$$b_3 = \sigma_1 + \mu_4V^2 + \tau_{11}V^4$$
$$b_5 = \mu_1 + \tau_4V^2$$
$$b_7 = \tau_1 \tag{2.5.7}$$

$$c_2 = 2[\sigma_2V + \mu_8V^3 + \tau_{16}V^5]$$
$$c_4 = 2[\mu_3V + (\tau_8-4\tau_{10})V^3]$$
$$c_6 = 2\tau_3V$$

$$d_3 = \mu_6 V^2 + \tau_{12} V^4$$

$$d_5 = \tau_6 V^2$$

$$e_4 = 8\tau_{10} V^3$$

$$\overline{b}_1 = (\sigma_3+\sigma_4) V^2 + \mu_{11} V^4 + \tau_{19} V^6$$

$$\overline{b}_3 = \sigma_1 + \mu_5 V^2 + \tau_{13} V^4$$

$$\overline{b}_5 = \mu_1 + \tau_5 V^2$$

$$\overline{b}_7 = \tau_1 \qquad (2.5.8)$$

$$\overline{c}_2 = 2[\sigma_2 V + \mu_9 V^3 + \tau_{17} V^5]$$

$$\overline{c}_4 = 2[\mu_3 V + (\tau_9 - 2\tau_{10}) V^3]$$

$$\overline{c}_6 = 2\tau_3 V$$

$$\overline{d}_3 = \mu_6 V^2 + \tau_{14} V^4$$

$$\overline{d}_5 = \tau_6 V^2$$

$$\overline{e}_4 = e_4$$

To carry out the integrations necessary for evaluating the quantities P_1, P_2 and P_3 it is necessary to know the shape of the entrance pupil. The previous section has been concerned with this problem. King has shown that the errors involved in approximating the pupil by an ellipse are quite small if a best least-squares fit is used. Having fitted an ellipse to the pupil it is then convenient to transform coordinates so that the aperture is reduced to a circle. Let P be a point in the aperture having coordinates (y,z) in the original coordinate system and (Y,Z) in the transformed coordinate system. Let the equation for the fitted ellipse be

$$\left(\frac{y-c}{a}\right)^2 + (z/b)^2 = 1 \qquad (2.5.9)$$

Then the transform relations are

$$Y = (y-c)/a, \qquad Z = z/b \qquad (2.5.10)$$

Let $\alpha = Y^2 + Z^2$, $\phi = \arctan(Z/Y)$

The aberration expansion in the transformed coordinates can be written

$$\begin{aligned} \varepsilon_y &= A^* + B^*\cos\phi + C^*\cos^2\phi + \ldots\ldots + H^*\cos^7\phi \\ \varepsilon_z &= \overline{B}^*\sin\phi + \overline{C}^*\cos\phi\sin\phi + \ldots\ldots + \overline{H}^*\cos^6\phi\sin\phi \end{aligned} \qquad (2.5.11)$$

As before let $X^* = \sum_{i=0}^{7} x_i^*\alpha^i$, for $X^* = A^*, B^*, \ldots \bar{H}^*$. Then

$$a_0^* = a_0+c.b_1+c^2.c_2+c^3.d_3+c^4.e_4+(a_2+c.b_3+c^2.c_4+c^3.d_5).c^2$$
$$+ (a_4+c.b_5+ c^2.c_6).c^4+(a_6+c.b_7).c^6;$$

$$a_2^* = (a_2+c.b_3+c^2.c_4+c^3.d_5).b^2+(a_4+c.b_5+c^2.c_6).2b^2c^2+(a_6+c.b_7).3b^2c^4;$$

$$a_4^* = (a_4+c.b_5+c^2.c^6).b^4+(a_6+c.b_7).3b^4c^2;$$

$$a_6^* = (a_6+c\ b_7).b^6;$$

$$b_1^* = (b_1+b_3c^2+b_5c^4+b_7c^6).a+(c_2+c_4c^2+c_6c^4).2ac+(d_3+d_5c^2).3ac^2+e_4 4ac^3$$
$$+(a_2+2a_4c^2+3a_6c^4).2ac+(b_3+2b_5c^2+3b_7c^4).2ac^2+(c_4+2c_6c^2).2ac^3+2d_5ac^4;$$

$$b_3^* = (b_3+2b_5c^2+ 3b_7c^4).ab^2+(c_4+2c_6c^2).2ab^2c+3d_5.ab^2c^2+4a_4.ab^2c$$
$$+12a_6.ab^2c^3+(b_5+3b_7.c^2).4ab^2c^2+c_6.4ab^2c^3;$$

$$b_5^* = (b_5+2c_6c).ab^4+3b_7.ab^4c^2+(a_6+b_7c).6ab^4c;$$

$$b_7^* = b_7.ab^6;$$

Let $\chi = b^2 - a^2$

$$c_2^* = (a_2+b_3.c+c_4.c^2+d_5.c^3)\chi+(b_3+2c_4.c+3d_5.c^2).2a^2c+(c_4+3d_5.c).a^2c^2$$
$$+ (a_4+b_5.c+c_6.c^2).2c^2(2a^2+\chi)+(b_5+2c_6.c).4a^2c^3+c_6.a^2c^4$$
$$+ (a_6+b_7.c).3c^4(4a^2+\chi)+6b_7.a^2c^5+c_2.a^2+3d_3.a^2c+6e_4.a^2c^2;$$

$$c_4^* = (a_4+b_5.c+c_6.c^2).2\chi b^2+(b_5+2c_6.c).4a^2b^2c+2c_6.a^2b^2c^2$$
$$+ (a_6+b_7.c).6b^2c^2(2a^2+\chi)+b_7.12a^2b^2c^3+c_4.a^2b^2+3d_5.a^2b^2c;$$

$$c_6^* = (a_6+cb_7).3\chi b^4+6b_7.a^2b^4c+c_6.a^2b^4;$$

$$d_3^* = (a_6+b_7.c).4ac^3(3\chi+2a^2)+3b_7.(4a^2+\chi).ac^4+(a_4+b_5.c+c_6.c^2).4a\chi c$$
$$+ (b_5+2c_6.c).2ac^2(2a^2+\chi)+4c_6.a^3c^3+(b_3+2c_4.c+3d_5.c^2).a\chi +(c_4+3d_5.c).2a^3c$$
$$+ d_5.a^3c^2+d_3.a^3+4e_4.a^3c;$$

$$d_5^* = (a_6+b_7.c).12ab^2\chi c+6b_7.ab^2c^2(2a^2+\chi)+(b_5+2c_6.c).2\chi ab^2+4c_6.a^3b^2c+d_5.a^3b^2;$$

$$d_7^* = 3b_7\chi ab^4;$$

$$e_4^* = (a_6+b_7.c).3\chi c^2(4a^2+\chi)+4b_7.a^2c^3(3\chi+2a^2)+(a_4+b_5.c+c_6.c^2).\chi^2+(b_5+2c_6.c)$$
$$.4\chi a^2c+2c_6.a^2c^2(2a^2+\chi)+(c_4+3d_5.c)a^2\chi+2d_5.a^4c+e_4.a^4;$$

$$e_6^* = (a_6+b_7.c).3\chi^2b^2+12b_7.\chi a^2b^2c+2c_6.\chi a^2b^2;$$

$$f_5^* = (a_6+b_7.c).6a\chi^2c+b_5.\chi^2a+3b_7.ac^2(\chi+4a^2)+4c_6.a^3\chi c+2c_6.\chi^2ac+d_5.\chi a^3;$$

$$f_7^* = 3b_7.\chi^2ab^2;$$

$g_6^* = (a_6+b_7\cdot c)\cdot\chi^3+6b_7a^2\chi^2c+c_6\cdot\chi^2a^2;$

$h_7^* = b_7\cdot\chi^3a;$

$\bar{b}_1^* = b\cdot[\bar{b}_1+\bar{c}_2\cdot c+\bar{d}_3\cdot c^2+\bar{e}_4\cdot c^3+(\bar{b}_3+\bar{c}_4\cdot c+\bar{d}_5\cdot c^2)\cdot c^2+(\bar{b}_5+\bar{c}_6\cdot c)\cdot c^4+\bar{b}_7\cdot c^6];$

$\bar{b}_3^* = b^3\cdot[\bar{b}_3+\bar{c}_4\cdot c+\bar{d}_5\cdot c^2+(\bar{b}_5+\bar{c}_6\cdot c)\cdot\ 2c^2+3\bar{b}_7\cdot c^4];$

$\bar{b}_5^* = b^5\cdot[\bar{b}_5+\bar{c}_6\cdot c+3\bar{b}_7\cdot c^2];$

$\bar{b}_7^* = b^7\cdot\bar{b}_7;$

$\bar{c}_2^* = ab\cdot[6\bar{b}_7\cdot c^5+(\bar{b}_5+\bar{c}_6\cdot c)\cdot 4c^3+\bar{c}_6\cdot c^4+(\bar{b}_3+\bar{c}_4\cdot c+\bar{d}_5\cdot c^2)\cdot 2c+(\bar{c}_4+2\bar{d}_5\cdot c)\cdot c^2$

$\quad +\bar{c}_2+2\bar{d}_3\cdot c+3\bar{e}_4\cdot c^2];$

$\bar{c}_4^* = ab^3\cdot|12\bar{b}_7\cdot c^3+(\bar{b}_5+\bar{c}_6\cdot c)\cdot 4c+2\bar{c}_6\cdot c^2+(\bar{c}_4+2\bar{d}_5\cdot c)|;$

$\bar{c}_6^* = ab^5\cdot(6\bar{b}_7\cdot c+\bar{c}_6);$

$\bar{d}_3^* = b\cdot[3\bar{b}_7\cdot(4a^2+\chi)\cdot c^4+(\bar{b}_5+\bar{c}_6\cdot c)\cdot 2c^2(2a^2+\chi)+4\bar{c}_6\cdot a^2c^3+(\bar{b}_3+\bar{c}_4\cdot c+\bar{d}_5\cdot c^2)\chi$

$\quad +(\bar{c}_4+2\bar{d}_5\cdot c)\cdot 2a^2c+\bar{d}_5\cdot a^2c^2+\bar{d}_3\cdot a^2+3\bar{e}_4\cdot a^2c];$

$\bar{d}_5^* = b^3\cdot[6\bar{b}_7\cdot(2a^2+\chi)c^2+(\bar{b}_5+\bar{c}_6\cdot c)\cdot 2\chi+4\bar{c}_6\cdot a^2c+\bar{d}_5a^2];$

$\bar{d}_7^* = 3\bar{b}_7\cdot\chi b^5;$

$\bar{e}_4^* = ab\cdot[\bar{b}_7\cdot 4c^3(3\chi+2a^2)+(\bar{b}_5+\bar{c}_6\cdot c)\cdot 4\chi c+2\bar{c}_6\cdot(2a^2+\chi)c^2+(\bar{c}_4+2\bar{d}_5\cdot c)\chi$

$\quad +2\bar{d}_5\cdot a^2c+\bar{e}_4\cdot a^2];$

$\bar{e}_6^* = 2ab^3\chi\cdot(6\bar{b}_7\cdot c+\bar{c}_6);$

$\bar{f}_5^* = b\chi\cdot[\bar{b}_5\cdot\chi+3\bar{b}_7\cdot(4a^2+\chi)c^2+4\bar{c}_6\cdot a^2c+\bar{c}_6\cdot\chi c+\bar{d}_5\cdot a^2];$

$\bar{f}_7^* = 3\bar{b}_7\cdot\chi^2b^3;$

$\bar{g}_6^* = ab^2\chi\cdot(6\bar{b}_7\cdot c+\bar{c}_6);$

$\bar{h}_7^* = \bar{b}_7\cdot b\chi^3;$ (2.5.12)

Using these expressions the three functions may be easily found as follows:

(a) Expression for P_3:

$$P_3 = \pi^{-1}\cdot\int_0^1\int_0^{2\pi}\varepsilon_y(\alpha,\phi)\alpha d\alpha d\phi$$

$$= \int_0^1\left[\int_0^{2\pi}\sum_{i=0}^{7} Z_i\ \cos^i\phi\ d\phi\right]\ \alpha d\alpha$$

where $Z_0 = A^*$, $Z_1 = B^*$, etc..

Hence

$$P_3 = a_0^*+(2a_2^*+c_2^*)/4+(2a_4^*+c_4^*+\tfrac{3}{8}e_4^*)/6+(2a_6^*+c_6^*+\tfrac{3}{4}e_6^*-\tfrac{5}{8}g_6^*)/8 \qquad (2.5.13)$$

(b) Expression for $A^{-1}\cdot\iint\varepsilon_y^2dA$:

Squaring the expression for ε_y in (3.8.7) then integrating with respect to ϕ gives

$$\pi^{-1}.\int_o^{2\pi}\varepsilon y^2 d\phi = 2A^{*2}+B^{*2}+2A^*C^*+\tfrac{3}{4}.[C^{*2}+2.(B^*D^*+A^*E^*)]+\tfrac{5}{8}[D^{*2}+2(A^*G^*+B^*F^*+C^*E^*)]+\tfrac{35}{64}[E^{*2}+2.(B^*H^*+C^*G^*+D^*F^*)]+\tfrac{63}{128}.(F^{*2}+2(D^*H^*+E^*G^*)]+\tfrac{231}{512}(G^{*2}+2F^*H^*)+\tfrac{429}{1024}H^{*2};$$

Grouping terms of equal power in α then integrating with respect to α gives

$$\pi^{-1}.\int_o^1\int_o^{2\pi}\varepsilon y^2\alpha d\alpha d\phi = \sum_{j=o}^{7} Y_{2j}/(2j+2) \qquad (2.5.14)$$

Expressions for the Yj terms are contained in the listing of program P5. A similar expression can be derived for $A^{-1}.\int\int\varepsilon_z{}^2 dA$, with Zj terms replacing Yj terms

$$\pi^{-1}.\int_0^1\int_0^{2\pi}\varepsilon_z{}^2\alpha d\alpha d\phi = \sum_{j=o}^{7} Z_{2j}/(2j+2) \qquad (2.5.15)$$

Using (2.5.14) and (2.5.15) expressions for P_1 and P_2, evaluated in the gaussian image plane, may be found.

§2.5(c) Non-Gaussian Image Planes

Let the chosen image plane be a distance x' from the gaussian image plane. Let $(\varepsilon_y{}',\varepsilon_z{}')$ denote the aberration, referred to an origin at the intersection point of the paraxial principal ray with the new image plane, of a ray (ρ,θ).

Then
$$\left.\begin{aligned}\varepsilon_y{}' &= \varepsilon_y + x'\ v'_{ak}\ \rho\cos\theta + x'.0(3)\\ \varepsilon_z{}' &= \varepsilon_z + x'\ v'_{ak}\ \rho\sin\theta + x'.0(3)\end{aligned}\right\} \qquad (2.5.16)$$

Suppose the aberration is given to a sufficient degree of accuracy by

$$\varepsilon_y{}' = \varepsilon_y + x'v'_{ak}\ \rho\cos\theta,\quad \varepsilon_z{}' = \varepsilon_z + x'v'_{ak}\ \rho\sin\theta \qquad (2.5.17)$$

Now in the transformed coordinates this becomes

$$\varepsilon_y{}' = \varepsilon_y + x'.\ v'_{ak}(a\alpha\cos\phi+c),\quad \varepsilon_z{}' = \varepsilon_z + x'\ v'_{ak}.b\alpha\sin\theta \qquad (2.5.18)$$

Then $\overline{\varepsilon}' = \pi^{-1}.\int_0^1\int_0^{2\pi}[\varepsilon_y + x'\ v'_{ak}(a\alpha\cos\phi+c)]\alpha d\alpha d\phi$

so $P_3' = P_3 + c\ x'v'_{ak}$ (2.5.19)

Also $P_1' = \pi^{-1}.\int_o^1\int_o^{2\pi}(\varepsilon_y{}'^2+\varepsilon_z{}'^2)\alpha d\alpha d\phi$

so $P_1' = P_1 + 2x'\ v'_{ak}.(aI_1+bI_2+c.P_3) + x'^2\ v'^2_{ak}\ [\tfrac{1}{4}(a^2+b^2)+c^2]$, (2.5.20)

where $I_1 = \pi^{-1}.\int_o^1\int_o^{2\pi}\varepsilon_y\alpha^2\cos\phi d\alpha d\phi$, $I_2 = \pi^{-1}.\int_o^1\int_o^{2\pi}\varepsilon_z\alpha^2\sin\phi\ d\alpha d\phi$, (2.5.21)

and $P_2' = P_2 + 2x'\ v'_{ak}(aI_1 - bI_2 + cP_3) + x'^2\ v'^2_{ak}[\tfrac{1}{4}(a^2-b^2)+c^2]$, (2.5.22)

If the centroid of the spot diagram, $(-P_3\ ,0)$, is chosen as origin for the measurement of the aberrations the expressions (3.9.5), (3.9.7) are simplified to become

$$P_{1c}' = P_{1c} + 2x'\ v_{ak}'(aI_1+bI_2) + \tfrac{1}{4}\ x'^2\ v_{ak}'^2\ (a^2+b^2) \qquad (2.5.23)$$

$$P_{2c}' = P_{2c} + 2x'\ v_{ak}'\ (aI_1-bI_2) + \tfrac{1}{4}x'^2\ v_{ak}'^2\ (a^2-b^2) \qquad (2.5.24)$$

the subscript c denoting that the quantity is calculated for aberrations measured from the centroid of the spot diagram.

§2.5(d) Calculation of O.T.F.-based Quantities

The o.t.f. is the normalised fourier transform of the diffraction image of a point object, where the normalising factor is such as to give unity for the zero frequency. As such it can be defined for any image forming system and specified object point irrespective of requirements of linearity and stationarity normally discussed in the introduction of the concept of the transfer function. However the transfer function becomes a valuable concept in discussing the performance of optical systems when these conditions are approximately satisfied. The literature on these is quite extensive and an excellent introduction is given in Goodman's monograph.[28]

Hopkins[29] has shown that the o.t.f. of an invariant, linear optical system can be expressed by

$$T(s,t) = A^{-1}_{(o,o)} \cdot \iint_{A(s,t)} \exp\{jk\ \Delta(s,t)\} dy\ dz \qquad (2.5.25)$$

where $A(s,t)$ denotes the region common to two sheared pupils with centres at $(\pm s, \pm t)$, and $\Delta(s,t)$ is the wavefront aberration difference function. The integral involved here cannot be evaluated analytically except under very restricted conditions, and so must be calculated numerically. A mesh summation method due to Hopkins[30] has been used for this work. In this method the integration region is divided into N^2 rectangular meshes and the contribution from the mesh element (p,q) is approximated as follows:

The mesh element (p,q) is defined by $y = y_p \mp \delta y$, $z = z_q \mp \delta z$. Approximate $\Delta(s,t)$ by a three term Taylor series,

$$\Delta(s,t) = \Delta(y,z,s,t) \simeq \Delta(y_p,z_q,s,t)+(y-y_p)\left(\frac{\partial\Delta}{\partial y}\right)_{y=y_p}+(z-z_p)\left(\frac{\partial\Delta}{\partial z}\right)_{z=z_q}$$

Then
$$I_{pq} = \int_{y_p-\delta_y}^{y_p+\delta_y}\int_{z_q-\delta_z}^{z_q+\delta_z} \exp\{jk\Delta\}dydz$$

$$= \exp\{jk\Delta(y_p,z_q,s,t)\}.\frac{\sin[\delta_y.\Delta_y(y_p,z_p,s,t)\ .\sin\ \delta_z.\Delta_z(y_p,z_q,s,t)]}{\delta_y.\Delta_y(y_p,z_q,s,t).\delta_z.\Delta_z(y_p,z_q,s,t)}$$

The I_{pq} are the approximate mesh element contributions.

The o.t.f. computation was programmed to allow o.t.f. values for both sagittal and tangential orientations to be calculated. The number of mesh elements, N^2, is specified by the user and a value of $N = 12$ was found quite satisfactory. The computed values were checked against those calculated by the S.I.R.A. Institute with agreement to within .02 MTF units in all cases. This provided a limited check on the accuracy of the retardation coefficients also. The calculation of the quantities S_1, S_2 and S_3 is, of course, trivial once the real and imaginary parts of the o.t.f. have been found.

§2.5(e) Summary

This section has given details of the expressions involved in calculating the spot diagram quantities and in the o.t.f. calculation. In §2.5(a) highly efficient means for determining entrance and exit pupils have been proposed. Using the aberration expansion (2.5.5) it has been shown that, after approximating the entrance pupil by an ellipse, the three spot diagram quantities can be readily calculated using the expressions (2.5.7), (2.5.8), (2.5.12 - 15). Furthermore the effect of choosing non-gaussian image planes and referring the aberrations to the centroid of the spot diagram has been studied and expressions (2.5.19 - 24) are the results of these considerations. Hopkins' method of calculating the o.t.f. has been summarized and so, using the defining relations (2.3.2), the new spatial frequency functions may be found.

§2.6 Some Additional Quantities

Besides the spatial frequency functions S_1, S_2, S_3 and the spot diagram quantities P_1, P_2, P_3 two other sets of assessment functions have been used for the comparison. These are the variance of the wavefront aberration, and the variance of the wavefront aberration difference function $\Delta(r,\psi)$ defined by

$$\Delta(r,\psi) = W(y+s,\ z+t) - W(y-s,\ z-t) \tag{2.6.1}$$

where $s = r \sin\psi, \quad t = r\cos\psi.$

The variance is considered for the two orthogonal sections $\psi = 0, \frac{\pi}{2}$. Obvious analogues to S_1 and S_2 can then be defined using these quantities and these will be considered during the comparison using actual systems.

§2.6(a) Wavefront Variance

The relationship between the wavefront variance and the Strehl ratio, together with the work of de Velis mentioned earlier, suggest the use of this quantity in a comparative study of assessment functions. Marechal's treatment of aberration tolerances for well corrected systems is based on obtaining a minimal value of 0.8 for the Strehl ratio, this requiring the variance of the wavefront aberration to be less than $\lambda^2/196$.

Suppose the wavefront aberration is written

$$W(\rho,\theta) = P_1(\rho,\theta) + P_2(\rho,\theta)$$

where P_1 is an even function of $\cos\theta$, P_2 is odd in $\cos\theta$. Then

$$V = \overline{W^2} - \overline{W}^2$$

so $$V = \overline{P_2{}^2} + \overline{P_1{}^2} - \overline{P_1}^2 \tag{2.6.2}$$

Let $$\begin{aligned} W(y,z,h) = {} & z_1\lambda+z_2y+\pi_1\lambda^2+\pi_2\lambda\mu+\pi_3\lambda\nu+\pi_4\mu^2+\pi_5\mu\nu+\sigma_1\lambda^3+\sigma_2\lambda^2\mu+\sigma_3\lambda^2\nu \\ & +\sigma_4\lambda\mu^2+\sigma_5\lambda\mu\nu+\sigma_6\lambda\nu^2+\sigma_7\mu^3+\sigma_8\mu^2\nu+\sigma_9\mu\nu^2+\tau_1\lambda^4+\tau_2\lambda^3\mu+\tau_3\lambda^3\nu \\ & +\tau_4\lambda^2\mu^2+\tau_5\lambda^2\mu\nu+\tau_6\lambda^2\nu^2+\tau_7\lambda\mu^3+\tau_8\lambda\mu^2\nu+\tau_9\lambda\mu\nu^2+\tau_{10}\lambda\nu^3+\tau_{11}\mu^4 \\ & +\tau_{12}\mu^3\nu+\tau_{13}\mu^2\nu^2+\tau_{14}\mu\nu^3+0(10) \end{aligned} \tag{2.6.3}$$

where $\lambda = y^2 + z^2$, $\mu = yh$, $\nu = h^2$. Then

$$\begin{aligned} P_1(\rho,\theta) = {} & a_{20}\rho^2+a_{40}\rho^4+a_{60}\rho^6+a_{80}\rho^8+a_{22}\rho^2\cos^2\theta+a_{42}\rho^4\cos^2\theta+a_{62}\rho^6\cos^2\theta \\ & +a_{44}\rho^4\cos^4\theta, \end{aligned} \tag{2.6.4}$$

$$P_2(\rho,\theta) = (b_{11}+b_{31}\rho^2+b_{51}\rho^4+b_{71}\rho^6+b_{33}\rho^2\cos^2\theta+b_{53}\rho^4\cos^2\theta).h\rho\cos\theta \tag{2.6.5}$$

$$
\begin{aligned}
\text{where}\quad a_{20} &= z_1+\pi_3\nu+\sigma^6\nu^2+\tau_{10}\nu^3 & b_{11} &= z_2+\pi_5\nu+\sigma_9\nu^2+\tau_{14}\nu^3 \\
a_{22} &= (\pi_4+\sigma_8\nu+\tau_{13}\nu^2).\nu & b_{31} &= \pi_2+\sigma_5\nu+\tau_9\nu^2 \\
a_{40} &= \pi_1+\sigma_3\nu+\tau_6\nu^2 & b_{33} &= (\sigma_7+\tau_{12}\nu).\nu \\
a_{42} &= (\sigma_4+\tau_8\nu).\nu & b_{51} &= \sigma_2+\tau_5\nu \\
a_{44} &= \tau_{11}\nu^2 & b_{53} &= \tau_7\nu \\
a_{60} &= \sigma_1+\tau_3\nu & b_{71} &= \tau_2 \\
a_{62} &= \tau_4\nu \\
a_{80} &= \tau_1
\end{aligned}
\tag{2.6.6}
$$

We then have

$$
\begin{aligned}
\overline{P_2{}^2} = {} & a^2h^2.\{\tfrac{1}{4}b_{11}{}^2+\tfrac{1}{3}b_{11}b_{31}a^2+\tfrac{1}{8}(b_{31}{}^2+2b_{11}b_{51})a^4+\tfrac{1}{5}(b_{11}b_{71}+b_{31}b_{51})a^6 \\
& +\tfrac{1}{2}(b_{51}{}^2+2b_{31}b_{71})a^8+\tfrac{1}{7}b_{51}b_{71}a^{10}+\tfrac{1}{16}.b_{71}{}^2a^{12}+\tfrac{1}{4}b_{11}b_{33}a^2 \\
& +\tfrac{3}{16}.(b_{11}b_{53}+b_{31}b_{33}).a^4+\tfrac{3}{20}.(b_{31}b_{53}+b_{51}b_{33})a^6+\tfrac{1}{8}(b_{51}b_{53}+b_{71}b_{33})a^8 \\
& +\tfrac{3}{28}.b_{71}b_{53}a^{10}+\tfrac{5}{64}.b_{33}{}^2a^4+\tfrac{1}{8}b_{33}b_{53}a^6+\tfrac{5}{96}.b_{53}{}^2a^8\}.
\end{aligned}
\tag{2.6.7}
$$

$$
\begin{aligned}
\overline{P_1{}^2}-\overline{P}_1{}^2 = {} & a^4.\{\tfrac{1}{12}.a_{20}{}^2+\tfrac{1}{6}.a_{20}a_{40}a^2+(\tfrac{4}{45}.a_{40}{}^2+\tfrac{3}{20}.a_{20}a_{60}).a^4 \\
& +(\tfrac{2}{15}.a_{20}a_{80}+\tfrac{1}{6}.a_{40}a_{60})a^6.+(\tfrac{9}{112}.a_{60}{}^2+\tfrac{16}{105}.a_{40}a_{80}).a^8+\tfrac{3}{20}.a_{60}a_{80}a^{10} \\
& +\tfrac{16}{225}.a_{80}{}^2a^{12}+\tfrac{1}{12}.a_{20}a_{22}+\tfrac{1}{12}.(a_{20}a_{42}+a_{22}a_{40})a^2+(\tfrac{3}{40}.a_{20}a_{62}+\tfrac{4}{45}.a_{40}a_{42} \\
& +\tfrac{3}{40}.a_{60}a_{22})a^4+(\tfrac{1}{12}.a_{40}a_{62}+\tfrac{1}{12}.a_{60}a_{42}+\tfrac{1}{15}.a_{80}a_{22}).a^6+(\tfrac{9}{112}.a_{60}a_{62} \\
& +\tfrac{8}{105}.a_{80}a_{42}).a^8+\tfrac{3}{40}.a_{80}a_{62}a^{10}+\tfrac{1}{16}.a_{22}{}^2+(\tfrac{1}{16}.a_{20}a_{44}+\tfrac{5}{48}.a_{22}a_{42}).a^2 \\
& +(\tfrac{17}{360}.a_{42}{}^2+\tfrac{7}{80}.a_{22}a_{62}+\tfrac{1}{15}.a_{40}a_{44})a^4+(\tfrac{1}{12}.a_{42}a_{62}+\tfrac{1}{16}.a_{60}a_{44})a^6 \\
& +(\tfrac{17}{448}.a_{62}{}^2+\tfrac{2}{35}.a_{80}a_{44}).a^8+\tfrac{3}{32}.a_{22}a_{44}.a^2+\tfrac{1}{12}.a_{42}a_{44}a^4 \\
& +\tfrac{7}{96}.a_{62}a_{44}a^6+\tfrac{3}{128}.a_{44}{}^2a^8\}.
\end{aligned}
\tag{2.6.8}
$$

Equations (2.6.4), (2.6.5) and (2.6.6) then allow rapid computation of the wavefront variance for a given optical system at any aperture and field angle for which the eighth order aberration expansion adequately represents the actual aberration.

§2.6(b) Variance of the Difference Function

The ratio of the response of a system at a specified frequency to what it would be for the same system in the absence of aberrations has been used as a criterion of image quality and for prescribing tolerances on various aberrations. Hopkins has shown that this ratio, $M(r,\psi)$, satisfies the inequality

$$M(r,\psi) \geq 1 - 2\pi^2/\lambda^2 . V(r,\psi)$$

where $V(r,\psi) = \overline{\Delta^2}(s,t) - \overline{\Delta(s,t)}^2$, $s = r \sin\psi$, $t = r\cos\psi$,

provided the right-hand side is positive. Empirical investigations have shown that, provided $M(s,\psi) \geq 0.7$ this variance, $V(r,\psi)$, provides a direct measure of image quality. Algebraic expressions have, therefore, been derived for $V(r,o)$ and $V(r,\frac{\pi}{2})$.

Let

$$\begin{aligned}
a_{11} &= z_1+\pi_3\nu+\sigma_6\nu^2+\tau_{10}\nu^3 & b_{11} &= z_2+\pi_5\nu+\sigma_9\nu^2+\tau_{14}\nu^3 \\
a_{31} &= \pi_1+\sigma_3\nu+\tau_6\nu^2 & b_{31} &= \pi_2+\sigma_5\nu+\tau_9\nu^2 \\
a_{32} &= \pi_4+\sigma_8\nu+\tau_{13}\nu^2 & b_{51} &= \sigma_2+\tau_5\nu \\
a_{51} &= \sigma_1+\tau_{3_\nu} & b_{52} &= \sigma_7+\tau_{12}\nu \qquad (2.6.10) \\
a_{52} &= \sigma_4+\tau_8\nu & b_{71} &= \tau_2 \\
a_{71} &= \tau_1, \quad a_{72} = \tau_4, & b_{72} &= \tau_7 \\
a_{73} &= \tau_{11}
\end{aligned}$$

(i) <u>Sagittal Case</u>

Let $e_{ij} = \rho^i \cos j\ \theta$, $\qquad E_{ij} = \iint_{A(s,t)} e_{ij}\, dA.$

Now write

$$\Delta(s,o) = P_1(\rho,\theta) + P_2(\rho,\theta)$$

where

$$P_1(\rho,\theta) = 4s.(A_{11}e_{11}+A_{31}e_{31}+A_{51}e_{51}+A_{71}e_{71}+A_{33}e_{33}+A_{53}e_{53}),$$

$$P_2(\rho,\theta) = 4s.(B_{22}+B_{42}\rho^2+B_{62}\rho^4+B_{44}\rho^2\cos^2\theta).\rho^2\sin\theta\ \cos\theta \qquad (2.6.12)$$

The A- and B- coefficients are given by

$$\begin{aligned}
A_{11} &= a_{11}+2a_{31}s^2+3a_{51}s^4+4a_{71}s^6, \\
A_{31} &= 2a_{31}+6a_{51}s^2+12a_{71}s^4+a_{52}\nu+2a_{72}\nu s^2 \\
A_{51} &= 2a_{72}\nu+3a_{51}+12a_{71}s^2 \qquad (2.6.13) \\
A_{71} &= 4a_{71} \\
A_{33} &= 4a_{51}s^2+16a_{71}s^4-a_{52}\nu-2a_{72}\nu s^2 \\
A_{53} &= 16a_{71}s^2-2a_{72}\nu
\end{aligned}$$

$$B_{22} = (b_{31}+2b_{51}s^2+3b_{71}s^4).h$$

$$B_{42} = (2b_{51}+b_{72}\nu+6b_{71}s^2).h$$

$$B_{62} = 3b_{71}.h \qquad (2.6.14)$$

$$B_{44} = (4b_{71}s^2-b_{72}\nu).h$$

Since the region of integration is symmetric,

$$\overline{\Delta(s,o)} = \overline{P}_1+\overline{P}_2 = 0$$

Hence $V(s,o) = \overline{P_1{}^2}+\overline{P_2{}^2}$

$$V(s,o) = 16A^{-1}(s,o).s^2\{A_{11}{}^2E_{22}+2A_{11}A_{31}E_{42}+(A_{31}{}^2+2A_{11}A_{51}).E_{62}+2A_{11}A_{71}E_{82}$$
$$+(A_{51}{}^2+2A_{31}A_{71}).E_{102}+2A_{31}A_{51}E_{82}+2A_{51}A_{71}E_{122}+A_{71}{}^2E_{142}+2A_{33}A_{11}E_{44}$$
$$+2(A_{33}A_{31}+A_{53}A_{11})E_{64}+2(A_{33}A_{51}+A_{53}A_{31})E_{84}+2(A_{33}A_{71}+A_{53}A_{51})E_{104}$$
$$+2A_{53}A_{71}A_{124}+2A_{33}A_{53}E_{86}+A_{53}{}^2E_{106}+A_{33}{}^2E_{66} +B_{22}{}^2(E_{42}-E_{44})$$
$$+2B_{22}B_{42}(E_{62}-E_{64})+B_{42}{}^2(E_{82}-E_{84})+B_{62}{}^2(E_{122}-E_{124})+2B_{22}B_{44}.(E_{64}-E_{66})$$
$$+2B_{42}B_{44}.(E_{84}-E_{86})+2B_{62}B_{44}.(E_{104}-E_{106})+B_{44}{}^2.(E_{86}-E_{88})\}.$$

(2.6.15)

(ii) <u>Tangential Case</u>

For this case it is convenient to define the -Y axis as $\theta = 0$, as shown, so that

$$y = -\rho\cos\theta, \qquad z = \rho\sin\theta$$

Now write $\Delta(t,\frac{\pi}{2}) = P_3(\rho,\theta) + P_4(\rho,\theta)$, (2.6.16)

where $P_3(\rho,\theta) = 2t.(C_{00}+C_{20}e_{20}+C_{40}e_{40}+C_{60}e_{60}+C_{22}e_{22}+C_{42}e_{42}+C_{62}e_{62}+C_{44}e_{44})$,

(2.6.17)

$$P_4(\rho,\theta) = 2t.(D_{11}e_{11}+D_{31}e_{31}+D_{51}e_{51}+D_{71}e_{71}+D_{33}e_{33}+D_{53}e_{53}) \qquad (2.6.18)$$

The C- and D- coefficients are as follows:

$$C_{00} = (b_{11}+b_{31}t^2+b_{51}t^4+b_{71}t^6+b_{52}\nu t^2+b_{72}\nu t^4).h$$

$$C_{20} = (b_{31}+2b_{51}t^2+3b_{71}t^4).h$$

$$C_{40} = (b_{51}+3b_{71}t^2).h$$

$$C_{60} = b_{71}.h \qquad (2.6.19)$$

$$C_{22} = (2b_{31}+8b_{51}t^2+18b_{71}t^4+3b_{52}\nu+9b_{72}\nu t^2).h$$

$$C_{42} = (4b_{51}+24b_{71}t_2+3b_{72}\nu).h$$

$$C_{62} = 6b_{71}.h$$

$$C_{44} = (8b_{71}t^2+2b_{72}\nu).h$$

$$D_{11} = 2.(a_{11}+2a_{31}t^2+3a_{51}t^4+4a_{71}t^6+a_{32}\nu+2a_{52}\nu t^2+3a_{72}\nu t^4+2a_{73}\nu^2 t^2)$$

$$D_{32} = 2.(2a_{31}+6a_{51}t^2+12a_{71}t^4+a_{52}\nu+4a_{72}\nu t^2)$$

$$D_{51} = 2.(3a_{51}+12a_{71}t^2+a_{72}\nu) \qquad (2.6.20)$$

$$D_{71} = 8a_{71}$$

$$D_{33} = 2.(4a_{51}t^2+a_{52}\nu+16a_{71}t^4+6a_{72}\nu t^2+2a_{73}\nu^2)$$

$$D_{53} = 2.(16a_{71}t^2+a_{72}\nu)$$

The variance is then given by

$$V(t,\tfrac{\pi}{2}) = \overline{P_3{}^2}+\overline{P_4{}^2}-\overline{P_3}^2 \qquad (2.6.21)$$

so

$$\begin{aligned} V(t,\tfrac{\pi}{2}) = {} & 4t^2.\bar{A}^{1}(o,t).\{C_{20}{}^2(E_{40}-E_{20}{}^2/E_{00})+2C_{20}C_{40}(E_{60}-E_{20}E_{40}/E_{00}) \\ & +C_{40}{}^2(E_{80}-E_{40}{}^2/E_{00})+2C_{40}C_{60}.(E_{100}-E_{40}E_{60}/E_{00})+C_{60}{}^2.(E_{120}-E_{60}{}^2/E_{00}) \\ & +2C_{20}.[C_{22}.(E_{42}-E_{20}E_{22}/E_{00})+C_{42}(E_{62}-E_{20}E_{42}/E_{00})+C_{62}(E_{82}-E_{20}E_{62}/E_{00}) \\ & +C_{44}(E_{64}-E_{20}E_{44}/E_{00})]+2C_{40}.[C_{22}(E_{62}-E_{40}E_{22}/E_{00})+C_{42}(E_{82}-E_{40}E_{42}/E_{00}) \\ & +C_{62}(E_{102}-E_{40}E_{62}/E_{00})+C_{44}(E_{84}-E_{40}E_{44}/E_{00})]+2C_{60}[C_{22}(E_{82}-E_{60}E_{22}/E_{00}) \\ & +C_{42}(E_{102}-E_{60}E_{42}/E_{00})+C_{62}(E_{122}-E_{60}E_{62}/E_{00})+C_{44}(E_{104}-E_{60}E_{44}/E_{00})] \\ & +C_{22}{}^2(E_{44}-E_{22}{}^2/E_{00})+2C_{22}C_{42}(E_{64}-E_{22}E_{42}/E_{00})+C_{42}{}^2(E_{84}-E_{42}{}^2/E_{00}) \\ & +2C_{22}C_{62}(E_{84}-E_{22}E_{62}/E_{00})+C_{62}{}^2(E_{124}-E_{62}{}^2/E_{00})+2C_{44}[C_{22}(E_{66}-E_{44}E_{22}/E_{00}) \\ & +C_{42}(E_{86}-E_{44}E_{42}/E_{00})+C_{62}(E_{106}-E_{44}E_{62}/E_{00})]+C_{44}{}^2(E_{88}-E_{44}{}^2/E_{00}) \\ & +D_{11}{}^2E_{22}+D_{31}{}^2E_{62}+D_{51}{}^2E_{102}+D_{71}{}^2E_{142}+2[D_{11}D_{31}E_{42}+D_{11}D_{51}E_{62} \\ & +(D_{11}D_{71}+D_{31}D_{51})E_{82}+D_{31}D_{71}E_{102}+D_{51}D_{71}E_{122}+D_{11}D_{33}E_{44}+(D_{31}D_{33} \\ & +D_{11}D_{53})E_{64}+(D_{51}D_{33}+D_{31}D_{53})E_{84}+(D_{71}D_{33}+D_{51}D_{53})E_{104}+D_{71}D_{53}E_{124}] \\ & +D_{33}{}^2E_{66}+2D_{33}D_{53}E_{86}+D_{53}{}^2E_{106}\} \end{aligned} \qquad (2.6.22)$$

§2.7 Conclusion

A range of assessment functions has been introduced and methods for their calculation given. Apart from the o.t.f. all quantities can be rapidly calculated for any field angle or aperture using the set of 28 ray or retardation coefficients. Use of paraxial defocussing terms allows consideration of image planes in the neighbourhood of the gaussian image plane conjugate to the chosen object plane. Relations which have been established between the s.d. quantities P_1, P_2 and P_3 and the o.t.f. quantities S_1, S_2 and S_3 justify the comparison of corresponding quantities and establish a link with the methods of assessment used by Dutton. The following chapter is concerned with detailed theoretical and numerical comparisons of these assessment functions for a range of optical systems.

REFERENCES

2.1 A.E. Conrady. Applied Optics and Optical Design (O.U.P., London 1929).

2.2 M. Herzberger, Modern Geometrical Optics (Interscience Publishers New York, 1958).

2.3 F.D. Cruickshank and G.A. Hills, J.O.S.A., 50, 379 (1960).

2.4 D.G. Hawkins and E.H. Linfoot, Roy. Astron. Soc. M.N., 105, 334 (1946).

2.5 F.A. Lucy, J.O.S.A., 46, 699 (1956).

2.6 A. Marechal, Rev. d'Optique, 26, 257 (1947).

2.7 B.R.A. Nijboer, Physica, 10, 679 (1943); 13, 605 (1947).

2.8 K. Nienhuis and B.R.A. Nijboer, Physica, 14, 590 (1948).

2.9 P.M. Duffieux, L'Intégrale de Fourier et ses Applications à L'Optique, Faculté des Sciences, Besançon (1946).

2.10 O.H. Schade, R.C.A. Rev., 9, 5 (pt.1), 245 (pt.2), 490 (pt.3), 653 (pt.4), (1948).

2.11 H.H. Hopkins, Proc. Phys. Soc. 70 B, 1162 (1957).

2.12 K. Miyamoto, J.O.S.A., 48, 57 (1958).

2.13 W.B. King and J. Kitchen, Appl. Opt., 7, 1193 (1968).

2.14 H.H. Hopkins, Proc. Phys. Soc., 70 B, 449 (1957).

2.15 E.H. Linfoot, Optica Acta, 5, 1 (1958).

2.16 F. Lanzl, Optical Information Processing (abstracts), Sydney, August 1974.

2.17 O.N. Stavroudis and L.E. Sutton, NBS monographs 93 (1965).

2.18 M. De, Optical Information Processing (abstracts), Sydney, August 1974.

2.19 D. Dutton, Appl. Opt., 11, 1091 (1972).

2.20 E. Wolf, J.O.S.A., 42, 547 (1952).

2.21 K. Miyamoto, J.O.S.A. 48, 57 (1958); 48, 567 (1958); 49, 35 (1959).

2.22 N.S. Kapany and J.J. Burke, J.O.S.A., 52, 1351 (1962).

2.23 J.B. De Velis, J.O.S.A., 55, 165 (1965).

2.24 Japan Optical Engineering Research Association, 7, 15 (1970).

2.25 W.B. King, Appl. Opt., 7, 489 (1968).

2.26 P.J. Sands, Thesis, Australian National University, (1967) ch. 10.

2.27 P.J. Sands, J.O.S.A., 63, 803 (1973).

2.28 J.W. Goodman, An Introduction to Fourier Optics (McGraw-Hill Book Co., San Francisco, 1968).

2.29 H.H. Hopkins, Proc. Roy. Soc. A, 217, 408 (1953).

2.30 H.H. Hopkins, Proc. Phys. Soc. B, 70, 1002 (1957).

CHAPTER 3

A THEORETICAL AND NUMERICAL COMPARISON OF SOME ASSESSMENT QUANTITIES

§3.1 Introduction

The use of different assessment functions in examining the state of correction of a set of optical systems has been found to give a different ranking of these systems in terms of imaging performance[1] and to predict different states of correction to be optimum[2]. The amount of such comparison work is, however, rather limited. This chapter, therefore, is directed towards giving an extended comparison of the assessment functions introduced previously. This is done through a comparative examination of the expressions derived assuming certain simple combinations of aberration terms, and also through a consideration of a wide range of actual systems with their much more complex aberration expansions. The use of aberration coefficients which are defined for a given object plane of the optical system, rather than object point, has allowed explicit consideration of the variation in the form of aberration balancing with field angle.

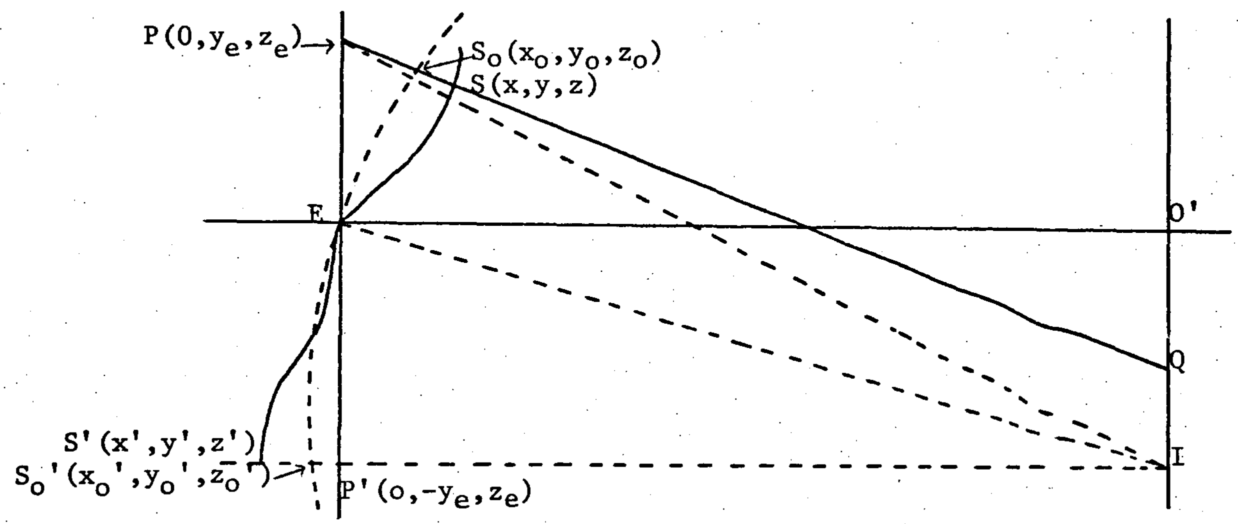
In the previous chapter expressions for the variance of the wavefront aberration and of the aberration difference function were derived, assuming a circular exit pupil. In fact, however, the integrations involved in calculating these quantities are over that region of the Y-Z plane corresponding to coordinates on the actual wavefront of transmitted rays. When the only vignetting aperture is a circular stop then the paraxial entrance and exit pupils are circular. However, for off-axis images, the region over which the integrations mentioned above must be carried is not circular. The next section of this chapter is therefore devoted to showing that it is possible to define a series of coordinate transformations which reduce the region of integration to a circle. These transformations can be described by modifications to the aberration coefficients which are field-independent, and the form of the retardation expansion is unaltered.

The following section gives a comparison of expressions obtained for optimum focal planes and for astigmatism and field curvature using various assessment quantities. A limited amount of comparative work on optimum focal planes has been done by De Velis[2] and his results are encompassed here.

The use of orthogonal expansions of the aberration function to deduce optimum forms of aberration balancing was introduced by Zernike[3]. This has been used to study optimum forms of balancing according to the minimization of wavefront variance[4,5] though such analyses generally avoid explicit studies of field-dependence. Sumita[6] has shown that the aberration difference function defined on the sheared pupil region can also be represented in terms of polynomials which are orthogonal on the unit circle, thus allowing a study of aberration balancing using this quantity. In section four it is shown that the spot diagram quantities $\overline{\varepsilon_y^{\,2}}$ and $\overline{\varepsilon_z^{\,2}}$ may also be expressed as the sums of squares of coefficients arising from the expansion of the transverse aberration function in terms of orthogonal polynomials. Expressions for these coefficients in terms of the third, fifth and seventh order $\overline{O}T$ coefficients are presented. Section five presents explicit expressions for the orthogonal coefficients in the expansion of the wavefront aberration using Zernike's circle polynomials, and section six does likewise for the sagittal and tangential difference function variance. These expressions are then compared, in section seven, for a number of simple cases, thus allowing a detailed algebraic study of optimum forms of aberration balancing using different criteria of image quality.

Finally we present in section eight some results from a numerical comparison of the various assessment functions introduced in chapter two. Using a range of optical designs having different degrees of correction the state of correction is examined using the spot diagram quantities, the variance of the wavefront retardation and of the difference function, and the functions defined in terms of o.t.f. values.

§3.2 Transformation of the Reference Sphere Pupil



In fig. 3.1 E′ is the axial point of the paraxial exit pupil and the origin of a rectangular coordinate system. O′ lies on the X-axis and is the axial point of the gaussian image plane. Let PQ be any ray from an object whose gaussian image is at I(e,-h,o). Let W_o be the gaussian reference sphere and W the aberrated wavefront. PQ intersects W_o and W at $S_o(x_o,y_o,z_o)$, $S(x,y,z)$ respectively. If $P(0,y_e,z_e)$ is any point on the periphery of the paraxial exit pupil then the reference sphere pupil periphery is defined by the set of points $S_o(x_o,y_o,z_o)$. The retardation function is defined by

$$R(y,z,h) = SS_o$$

and is written in a power series expansion as

$$R(y,z,h) = \pi_1\lambda^2+\pi_2\lambda\mu+\pi_3\lambda\nu+\pi_4\mu^2+\pi_5\mu\nu+\sigma_1\lambda^3+\sigma_2\lambda^2\mu+\sigma_3\lambda^2\nu$$
$$+\sigma_4\lambda\mu^2+\sigma_5\lambda\mu\nu+\sigma_6\lambda\nu^2+\sigma_7\mu^3+\sigma_8\mu^2\nu+\sigma_9\mu\nu^2$$
$$+\tau_1\lambda^4+\tau_2\lambda^3\mu+\tau_3\lambda^3\nu+\tau_4\lambda^2\mu^2+\tau_5\lambda^2\mu\nu+\tau_6\lambda^2\nu^2+\tau_7\lambda\mu^3$$
$$+\tau_8\lambda\mu^2\nu+\tau_9\lambda\mu\nu^2+\tau_{10}\lambda\nu^3+\tau_{11}\mu^4+\tau_{12}\mu^3\nu+\tau_{13}\mu^2\nu^2+\tau_{14}\mu\nu^3+0(10),$$

where $\lambda = y^2+z^2$, $\mu = yh$, $\nu = h^2$. If the retardation is now expressed in terms of coordinates (y_o,z_o) on the reference sphere then

$$y = y_o[1-e^{-1}(1+e^{-2}\nu)^{-\frac{1}{2}}R]-e^{-1}(1+e^{-2}\nu)^{-\frac{1}{2}}R.h+0(7) \tag{3.2.1a}$$

$$z = z_o[1-e^{-1}(1+e^{-2}\nu)^{-\frac{1}{2}}R]+0(7). \tag{3.2.1b}$$

Since R is 0(4) this change in coordinates will not affect either fourth or sixth order terms of the aberration expansion. Using primes to indicate the new coefficients, the modifications to the eighth order coefficients are as follows:-

$$\tau_1' = \tau_1 - 4e^{-1}\pi_1^2$$
$$\tau_2' = \tau_2 - e^{-1}(7\pi_1\pi_2 - 4\pi_1^2)$$
$$\tau_3' = \tau_3 - e^{-1}(6\pi_1\pi_3 - \pi_1\pi_2)$$
$$\tau_4' = \tau_4 - e^{-1}(6\pi_1\pi_4 - 6\pi_1\pi_2 + 3\pi_2^2)$$
$$\tau_5' = \tau_5 - e^{-1}(5\pi_1\pi_5 - 2\pi_1\pi_4 - 6\pi_1\pi_3 + 5\pi_2\pi_3 - \pi_2^2)$$
$$\tau_6' = \tau_6 - e^{-1}(2\pi_3^2 - \pi_1\pi_5 - \pi_2\pi_3)$$
$$\tau_7' = \tau_7 - e^{-1}(5\pi_2\pi_4 - 4\pi_1\pi_4 - 2\pi_2^2)$$
$$\tau_8' = \tau_8 - e^{-1}(4\pi_2\pi_5 + 4\pi_3\pi_4 - 4\pi_1\pi_5 - 4\pi_2\pi_3 - 3\pi_2\pi_4)$$
$$\tau_9' = \tau_9' - e^{-1}(3\pi_3\pi_5 - 2\pi_2\pi_5 - 2\pi_3^2 - 2\pi_3\pi_4)$$
$$\tau_{10}' = \tau_{10} + e^{-1}\pi_3\pi_5$$
$$\tau_{11}' = \tau_{11} + e^{-1}(2\pi_2\pi_4 - 2\pi_4^2)$$

$$\tau_{12}' = \tau_{12} + e^{-1}(2\pi_2\pi_5 + 2\pi_3\pi_4 + 2\pi_4^{\ 2} - 3\pi_4\pi_5)$$

$$\tau_{13}' = \tau_{13} + e^{-1}(2\pi_3\pi_5 + 3\pi_4\pi_5 - \pi_5^{\ 2})$$

$$\tau_{14}' = \tau_{14} + e^{-1}\pi_5^{\ 2} \qquad (3.2.2)$$

Consider now the location of the centre of the exit pupil. Suppose the paraxial exit pupil is circular and the lines joining the tangential extrema of the pupil to the gaussian image point intersect the reference sphere at $(y_o,0)$, $(y_o',0)$, - see fig. 1. Then the centre of the pupil is the point $(y_o+y_o')/2,0)$. By simple geometry we have

$$y_o = y_e(1 - x_o/e) + x_o h/e, \qquad z_o = z_e(1 - x_o/e) \qquad (3.2.3)$$

where $x_o/e = 1 - [1 - e^{-2}(\lambda_o - 2\mu_o)]^{\frac{1}{2}}$, $\lambda_o = y_o^{\ 2} + z_o^{\ 2}$, $\mu_o = y_o h$.

Using an $O(2)$ approximation for x_o gives

$$x_o/e = \tfrac{1}{2}\, e^{-2}(\rho_e^{\ 2} - 2\, y_e h)$$

where $\rho_e^{\ 2} = y_e^{\ 2} + z_e^{\ 2}$. Using this, a third order approximation to the tangential section extremities of the pupil is given by

$$y_{o\ max} = \rho - \tfrac{1}{2}e^{-2}(\rho^3 - 3\rho^2 h + 2\rho h^2) \qquad (3.2.4)$$

where $\rho = \mp$ maximum value of ρ_e. Hence a third order approximation to the centre of the pupil is

$$(y_o + y_o')/2 = 3/2\ e^{-2}\rho^2 h + O(5) \qquad (3.2.5)$$

In fig.3.2 the error, expressed as a fraction of full aperture, in finding the mid point of the tangential extrema using this approximation is shown as a function of aperture and semi-field angle. If a transformation of the origin of coordinates is made to this centre then the exit pupil (defined on the reference sphere) is symmetric in the new coordinates.

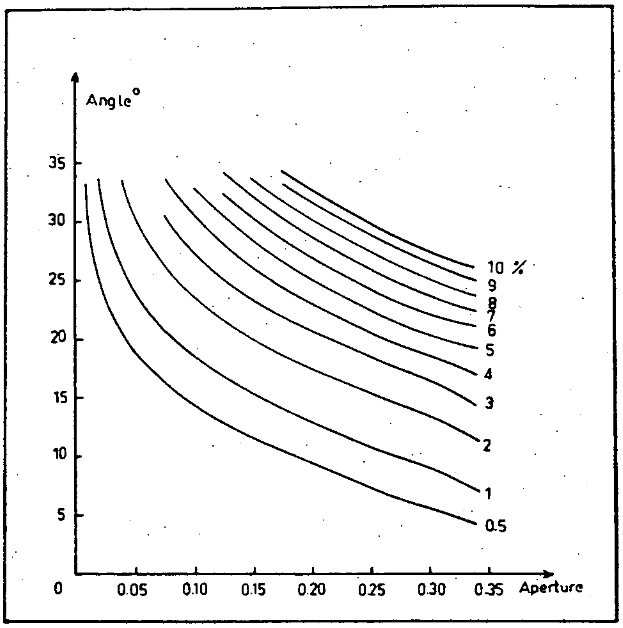


Fig. ~~2~~ 3·2 Error in location of centre.

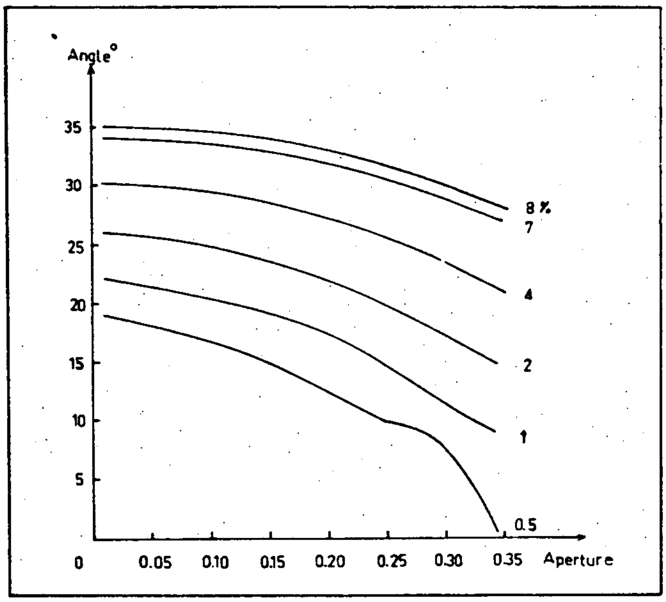


Fig. 3·3 Error of third order approximation to pupil values.

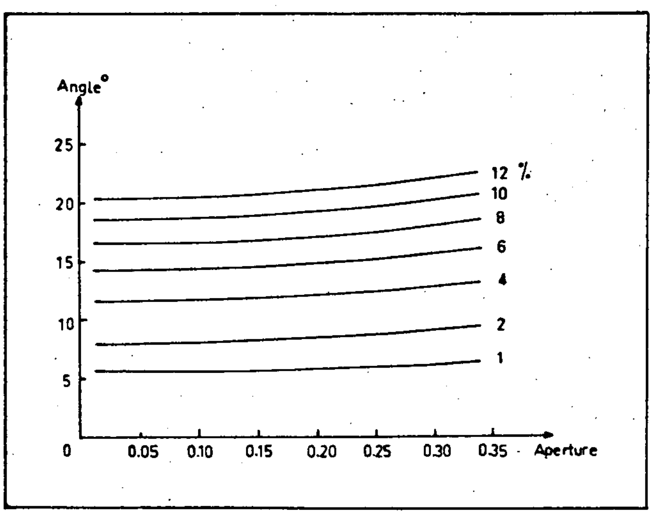


Fig. ~~X~~ 3·4 Deviation of pupil shape from circular.

The appropriate changes to the coefficients are given below. Primes indicate the transformed coefficients and $\alpha = 3/2e^{-2}\rho^2 = 3/8F^2$. Three new coefficients π_6', σ_{10}', and τ_{15}' are introduced, but these correspond simply to transverse defocussing and may be neglected.

$$\pi_1' = \pi_1, \quad \pi_2 = \pi_2' + 4\alpha\pi_1, \quad \pi_3' = \pi_3 + \alpha\pi_2 + 2\alpha^2\pi_1,$$
$$\pi_4' = \pi_4 + 2\alpha\pi_2 + 4\alpha^2\pi_1, \ \pi_5' = \pi_5 + 2\alpha(\pi_3 + \pi_4) + 3\alpha^2\pi_2 + 4\alpha^3\pi_1,$$
$$\pi_6' = \alpha\pi_5 + \alpha^2(\pi_3+\pi_4) + \alpha^3\pi_2 + \alpha^4\pi_1, \tag{3.2.6}$$

$$\sigma_1' = \sigma_1$$
$$\sigma_2' = \sigma_2 + 6\alpha\sigma_1$$
$$\sigma_3' = \sigma_3 + \alpha\sigma_2 + 3\alpha^2\sigma_1$$
$$\sigma_4' = \sigma_4 + 4\alpha\sigma_2 + 12\alpha^2\sigma_1$$
$$\sigma_5' = \sigma_5 + 2\alpha(\sigma_4 + 2\sigma_3) + 6\alpha^2\sigma_2 + 12\alpha^3\sigma_1$$
$$\sigma_6' = \sigma_6 + \alpha\sigma_5 + \alpha^2(\sigma_4 + 2\sigma_3) + 2\alpha^3\sigma_2 + 3\alpha^4\sigma_1$$
$$\sigma_7' = \sigma_7 + 2\alpha\sigma_4 + 4\alpha^2\sigma_2 + 8\alpha^3\sigma_1$$
$$\sigma_8' = \sigma_8 + \alpha(3\sigma_7+\sigma_5) + \alpha^2(5\sigma_4 + 4\sigma_3) + 8\alpha^3\sigma_2 + 12\alpha^4\sigma_1$$
$$\sigma_9' = \sigma_9 + 2\alpha(\sigma_6 + \sigma_8) + 3\alpha^2(\sigma_5 + \sigma_7) + 4\alpha^3(\sigma_3 + \sigma_4) + 5\alpha^4\sigma_2 + 6\alpha^5\sigma 1$$
$$\sigma_{10}' = \alpha\sigma_9 + \alpha^2(\sigma_6 + \sigma_8) + \alpha^3(\sigma_5 + \sigma_7) + \alpha^4(\sigma_3 + \sigma_4) + \alpha^5\sigma_2 + \alpha^6\sigma_1 \tag{3.2.7}$$

$$\tau_1' = \tau_1$$
$$\tau_2' = \tau_2 + 8\alpha\tau_1$$
$$\tau_3' = \tau_3 + \alpha\tau_2 + 4\alpha^2\tau_1$$
$$\tau_4' = \tau_4 + 6\alpha\tau_2 + 24\alpha^2\tau_1$$
$$\tau_5' = \tau_5 + 2\alpha(\tau_4 + 3\tau_3) + 9\alpha^2\tau_2 + 24\alpha^3\tau_1$$
$$\tau_6' = \tau_6 + \alpha\tau_5 + \alpha^2(\tau_4 + 3\tau_3) + 3\alpha^3\tau_2 + 6\alpha^4\tau_1$$
$$\tau_7' = \tau_7 + 4\alpha\tau_4 + 12\alpha^2\tau_2 + 32\alpha^3\tau_1$$
$$\tau_8' = \tau_8 + \alpha(3\tau_7 + 4\tau_5) + \alpha^2(10\tau_4 + 12\tau_3) + 24\alpha^3\tau_2 + 48\alpha^4\tau_1$$
$$\tau_9' = \tau_9 + 2\alpha(\tau_8 + 2\tau_6) + 3\alpha^2(\tau_7 + 2\tau_5) + 4\alpha^3(2\tau_4 + 3\tau_3) + 15\alpha^4\tau_2 + 24\,\alpha^5\tau_1$$

$$\tau_{10}' = \tau_{10} + \alpha\tau_9 + \alpha^2(\tau_8 + 2\tau_6) + \alpha^3(\tau_7 + 2\tau_5) + \alpha^4(2\tau_4 + 3\tau_3) + 3\alpha^5\tau_2 + 4\alpha^6\tau_1$$

$$\tau_{11}' = \tau_{11} + 2\alpha\tau_7 + 4\alpha^2\tau_4 + 8\alpha^3\tau_2 + 16\alpha^4\tau_1$$

$$\tau_{12}' = \tau_{12} + 4\alpha\tau_{11} + 2\alpha\tau_8 + \alpha^2(7\tau_7 + 4\tau_5) + 4\alpha^3(3\tau_4 + 2\tau_3) + 20\alpha^4\tau_2 + 32\alpha^5\tau_1$$

$$\tau_{13}' = \tau_{13} + 2\alpha\tau_9 + \alpha^2(6\tau_{11} + 5\tau_8 + 4\tau_6) + \alpha^3(9\tau_7 + 8\tau_5) + \alpha^4(13\tau_4 + 12\tau_3) + 18\alpha^5\tau_2 + 24\alpha^6\tau_1$$

$$\tau_{14}' = \tau_{14} + 2\alpha(\tau_{13} + \tau_{10}) + 3\alpha^2(\tau_{12} + \tau_9) + 4\alpha^3(\tau_{11} + \tau_8 + \tau_6) + 5\alpha^4(\tau_7 + \tau_5) + 6\alpha^5(\tau_4 + \tau_3) + 7\alpha^6\tau_2 + 8\alpha^7\tau_1$$

$$\tau_{15}' = \alpha\tau_{14} + \alpha^2(\tau_{13} + \tau_{10}) + \alpha^3(\tau_{12} + \tau_9) + \alpha^4(\tau_{11} + \tau_8 + \tau_6) + \alpha^5(\tau_7 + \tau_5) + \alpha^6(\tau_4 + \tau_3) + \alpha^7\tau_2 + \alpha^8\tau_1 \tag{3.2.8}$$

Two methods may be used to reduce the pupil to a circular region. The first is to average the third order tangential and sagittal extrema. This gives a third order approximation to the pupil radius, $\rho_o(3)$, where

$$\rho_o(3) = \rho_e[1 - \tfrac{1}{2}e^{-2}(\rho_e^2 + h^2)] \tag{3.2.9}$$

The accuracy of this approximation is shown in fig.3.3, where the fractional difference between the average of the actual extrema and the third order approximation is plotted as a function of aperture and angle. The second method, explained in the following, is far more satisfactory. In fig.3.4 the fractional difference between sagittal and tangential extrema of the exact reference sphere pupil is shown and it can be seen that the deviation of the pupil from being circular is almost independent of aperture. A good approximation to the shape of the pupil was found, numerically, to be

$$y_{max} = z_{max}(1 - c_1\nu) \tag{3.2.10}$$

where $c_1 \simeq 0.9$. This corresponds to the eccentricity of the ellipse being proportional to h, provided h is small. In fact c_1 decreases slowly with aperture ranging from 0.95 at F/5 to 0.82 at F/1.6. This could be considered by writing $c_1 = 1 - c_2\lambda$ and determining an appropriate value for c_2. However the aperture variation will not be considered in the following work. By now making the transformation

$$y' := y/(1 - c_1\nu) \tag{3.2.11}$$

the reference sphere pupil can be reduced to a circle. The effect of this last transformation on the coefficients is given below, with primes again indicating the transformed coefficients:-

$$\begin{aligned}
&\pi_j' = \pi_j, && j = 1, 2, 3, 4, 5, \\
&\sigma_j' = \sigma_j, && j = 1, 2, 3, 6, \\
&\sigma_4 = \sigma_4 - 4c_1\pi_1, \quad && \sigma_5 = \sigma_5 - c_1\pi_2, \quad \sigma_7 = \sigma_7 - 2c_1\pi_2 \\
&\sigma_8' = \sigma_8 - 2c_1(\pi_3 + \pi_4), \; \sigma_9' = \sigma_9 - c_1\pi_5.
\end{aligned} \tag{3.2.12}$$

$$\begin{aligned}
&\tau_j' = \tau_j, && j = 1, 2, 3, 6, 10 \\
&\tau_4' = \tau_4 - 6c_1\sigma_1, \quad && \tau_5' = \tau_5 - c_1\sigma_2, \quad \tau_7 = \tau_7 - 4c_1\sigma_2 \\
&\tau_8' = \tau_8 + 2c_1^2\pi_1 - 2c_1(2\sigma_3 + \sigma_4), && \tau_9' = \tau_9 - c_1\sigma_5 \\
&\tau_{11}' = \tau_{11} + 4c_1^2\pi_1, && \tau_{12}' = \tau_{12} + c_1^2\pi_2 - 3c_1\sigma_7 \\
&\tau_{13}' = \tau_{13} + c_1^2(\pi_3 + \pi_4) - 2c_1(\sigma_6 + \sigma_8), \; \tau_{14}' = \tau_{14} - c_1\sigma_9
\end{aligned} \tag{3.2.13}$$

The three sets of coefficient alterations given in this section provide a means whereby the retardation coefficients may be used to describe the wavefront aberration defined in terms of coordinates on a circular exit pupil which is centred on the origin of coordinates. Then, by expressing the pupil coordinates in polar form, it is a simple matter to carry out the integrations involved in calculating quantities such as the variance of the wavefront aberration.

§3.3 Longitudinal Shift of Focus

When discussing the aberrations associated with non-gaussian image planes in the neighbourhood of the gaussian image plane it is common practice to add a single longitudinal defocussing term to the aberration expansion. By considering the deformation function introduced earlier it is a simple matter to determine the conditions under which this provides a sufficiently accurate representation of the aberrations in non-gaussian image planes.

Let $(x-e)^2 + (y+h)^2 + z^2 + 2eD = e^2+h^2,$ (3.3.1)

and $(x-e')^2 + (y+h')^2 + z^2 + 2e'D' = e'^2 + h'^2,$ (3.3.2)

be two equations for the aberrated wavefront. $P(x,y,z)$ is any point on the aberrated wavefront. $I(e,-h,0)$ is the gaussian image point, and $I'(e',-h',0)$ is the corresponding out-of-focus ideal image point where $e' = e+s$, $h' = h(1+s/e)$, s being the axial displacement of the non-gaussian image plane from the gaussian plane. Elementary algebra then gives

$$D' - D = \tfrac{1}{2}e^{-2}s\lambda + \text{terms in } s^2 \text{ and } sD \qquad (3.3.3)$$

Hence $R' - R = \tfrac{1}{2}e^{-2}s\lambda$ + terms in s^2 and sD, showing that when s is sufficiently small for sR or s^2 to be negligible compared to the aberration then the single defocussing term is adequate. Also,

since $\underline{\varepsilon} = e \dfrac{\partial D}{\partial \underline{y}}$

the transverse aberration defocussing expression

$$\underline{\varepsilon}' = \underline{\varepsilon} + e^{-1}s\underline{y} \tag{3.3.4}$$

will have analogous conditions on its accuracy.

We will now discuss optimum image plane shifts according to various criteria, assuming that the amount of defocussing is sufficiently small for (3.3.3,4) to give adequate representations of the effect of defocussing on the aberration expansions.

i) Transverse Aberrations

Sands[7] has derived an expression for the surface of best focus using minimization of P_1 (the radius of gyration of the spot diagram), and has discussed the discrepancies between the values obtained by raytrace and those obtained using the truncated aberration expansion. The only significant cause for concern when using the aberration expansion was shown to be the convergence of this expansion. The apertures and angles for which analyses are performed in this work all lie within the region of convergence of the third, fifth and seventh order series expansion for the aberration function.

Using the notation of §2.5 we have

$$\begin{aligned}\overline{\varepsilon_y'^2} &= \overline{\varepsilon_y^2} + 2x'v_{ak}'.\overline{\varepsilon_y\rho\cos\theta} + \tfrac{1}{4}(x'v_{ak}'a)^2\\ \overline{\varepsilon_z'^2} &= \overline{\varepsilon_z^2} + 2x'v_{ak}'.\overline{\varepsilon_z\rho\sin\theta} + \tfrac{1}{4}(x'v_{ak}'a)^2\end{aligned} \tag{3.3.5}$$

where a is the radius of the entrance pupil. Note that we are here assuming a circular entrance pupil - a valid assumption when only one component of the system acts as a stop. Using the expression of §2.5 and noting that a=b and c=o (i.e. the pupil is circular and distortion is negligible) expressions for those planes which minimize $\overline{\varepsilon_y'^2}$ and $\overline{\varepsilon_z'^2}$ can be found. From (3.3.5) it also follows that the optimum plane found by minimizing P_1 is located midway between the planes minimizing $\overline{\varepsilon_y'^2}$ and $\overline{\varepsilon_z'^2}$. Denote these planes by x_{Gopt},

x_{Yopt} and x_{Zopt} respectively. Then we have

$$x_{Gopt} = -v_{ak}'^{-1}.\{\tfrac{2}{3}p_1a^2+(2p_3+p_4)V^2+\tfrac{1}{2}s_1a^4+\tfrac{1}{3}(s_4+s_5+s_6)a^2V^2$$
$$+\tfrac{1}{2}(s_{10}+s_{11})V^4+\tfrac{2}{3}t_1a^6+\tfrac{1}{4}(t_4+t_5+t_6)a^4V^2+\tfrac{1}{3}(t_{11}+\tfrac{3}{4}t_{12}+t_{13}+\tfrac{1}{4}t_{14})$$
$$a^2V^4+\tfrac{1}{2}(t_{18}+t_{19})V^6\}, \qquad (3.3.6)$$

$$x_{Yopt} = -v_{ak}'^{-1}.\{(3p_3+p_4)V^2+s_{10}V^4+t_{18}V^6+\tfrac{2}{3}a^2[p_1+(s_4+\tfrac{3}{4}s_6)V^2+(t_{11}$$
$$+\tfrac{3}{4}t_{12})V^4]+\tfrac{1}{2}a^4[s_1+(t_4+\tfrac{3}{4}t_6)V^2]+\tfrac{2}{5}t_1a^6\}, \qquad (3.3.7)$$

$$x_{Zopt} = -v_{ak}'^{-1}.\{(p_3+p_4)V^2+s_{11}V^4+t_{19}V^6+\tfrac{2}{3}a^2[p_1+(s_5+\tfrac{1}{4}s_6)V^2+(t_{13}+\tfrac{1}{4}t_{14})V^4]$$
$$+\tfrac{1}{2}a^4[s_1+(t_5+\tfrac{1}{4}t_6)V^2]+\tfrac{2}{5}t_1a^6\}. \qquad (3.3.8)$$

Note that $v_{ak}' = (m\ell_{01})^{-1}$ where m is the paraxial magnification of the system and ℓ_{01} the object distance.

ii) Wavefront Variance

Using (3.3.3) the variance of the wavefront retardation may be written as a quadratic function of the displacement, x', from the gaussian image plane. From this the optimum image plane found using minimization of this variance is given by

$$x_{Vopt} = -e^2\{2\pi_1a^2+(2\pi_3+\pi_4)h^2+\tfrac{9}{5}\sigma_1a^4+(2\sigma_3+\sigma_4)a^2h^2+(2\sigma_6+\sigma_8)h^4$$
$$+\tfrac{8}{5}\tau_1a^6+\tfrac{9}{5}(\tau_3+\tau_4)a^4h^2+(2\tau_6+\tau_8+\tfrac{3}{4}\pi_{11})a^2h^4+(2\pi_{10}+\pi_{11})h^6\}.$$
$$(3.3.9)$$

iii) Optical Transfer Function

A common method of selecting the optimum image plane according to transfer function theory is to choose some spatial frequency and find that plane for

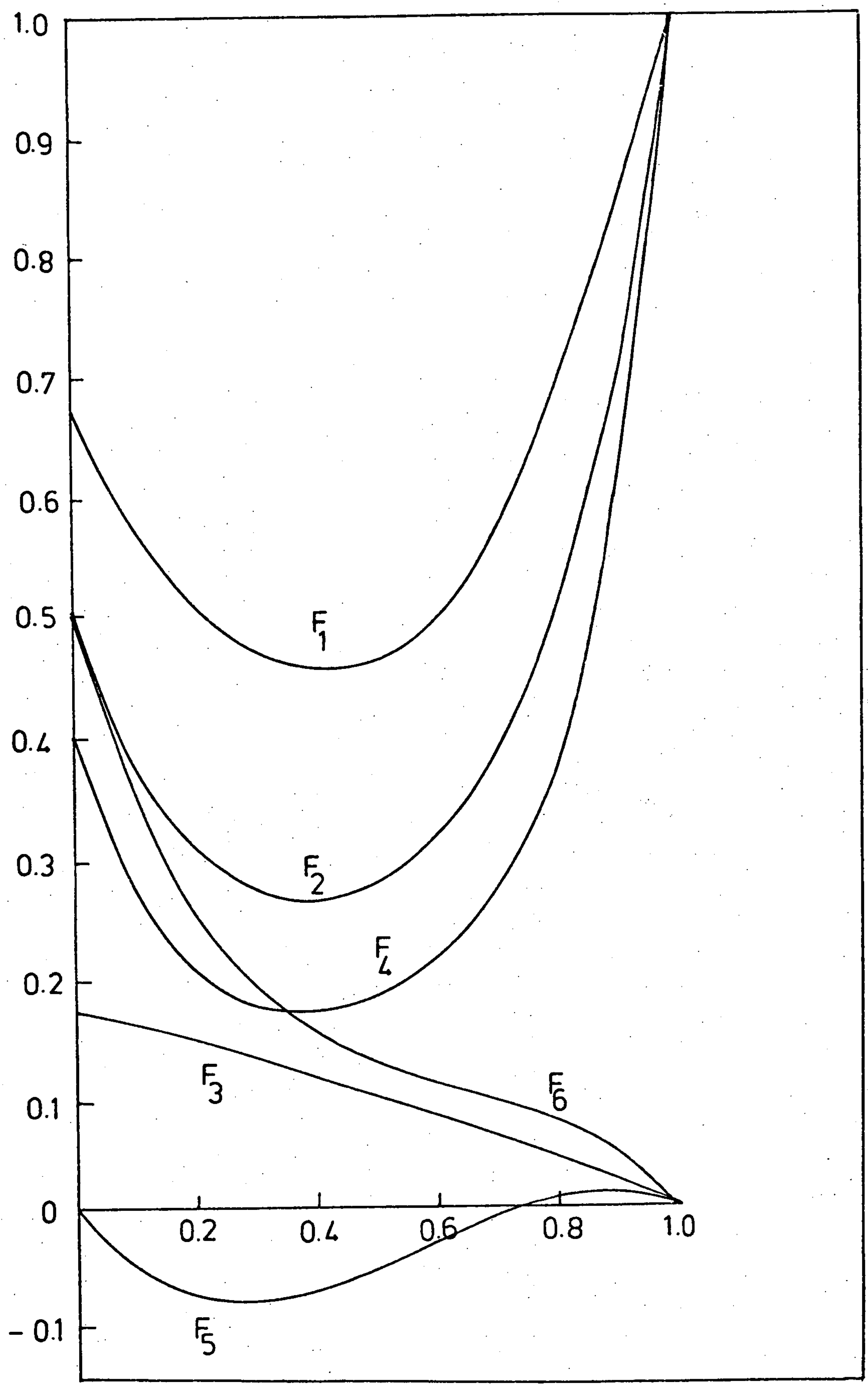


fig. 3.26

which the response at the chosen frequency is a maximum. Work by Fellgett and Linfoot[8] has shown that the information content of an optical image is concentrated in the higher frequency components of that image. Thus the spatial frequency chosen for such optimization ought to be near the upper limit of the frequency band of interest. Also, since the frequency response variation with image plane position has sharper maxima for higher frequencies (as illustrated in fig. 3.5), the relative effect of a displacement of the image plane from the optimum plane according to a high frequency is greater than that for a lower frequency.

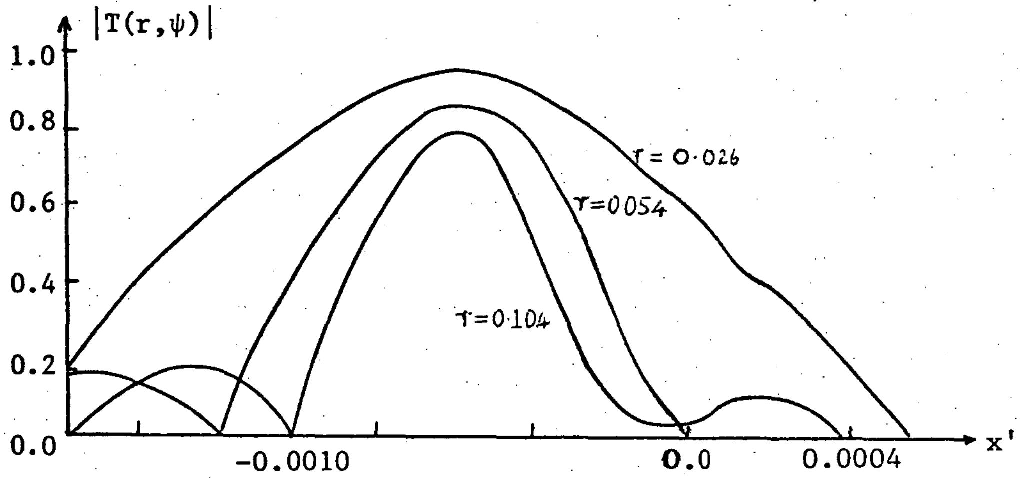


fig. 3.5

An analytic expression for the optimum image plane using the o.t.f. is not available and hence calculation of this is very slow. However such calculations have been done and results are provided in the comparisons using actual systems presented in §3.9.

iv) <u>Difference Function Variance</u>

Using the notation of §2.6(b) let us write $E_{ij} = E_{ij}/E_{22}$ and define functions F_j $(j = 1,2,3,4,5,6)$ by

$$F_1 = r^2 + E'_{42}$$

$$F_2 = r^4 + 2E'_{42}.r^2 + \frac{4}{3} E'_{44}\, r^2 + E'_{62}$$

$$F_3 = E'_{42} - E'_{44}$$

$$F_4 = r^6 + (3E'_{42} + 4E'_{44})\, r^4 + (3E'_{62} + 4E'_{64})\, r^2 + E'_{82}$$

$$F_5 = (E'_{42} - E'_{44})\, r^2 + (E'_{62} - E'_{64})$$

$$F_6 = E'_{62} + \frac{5}{2}\,(E'_{42} - E'_{44})\, r^2. \qquad (3.3.10)$$

The variance of the aberration difference function can be expressed as a quadratic function of the image plane shift x', and thus the value of x' which minimizes this variance can be found. Using the expressions of (2.6.15) and (2.6.22) optimum image planes have been found for sagittal and tangential line orientations. These are

$$x_{Sopt} = -2e^2\{2(\pi_1+\sigma_3h^2+\tau_6h^4)F_1+(\pi_3+\sigma_6h^2+\tau_{10}h^4)h^2+3(\sigma_1+\tau_3h^2)F_2$$
$$+(\sigma_4+\tau_8h^2)h^2.F_3+4\tau_1F_4+2\tau_4h^2F_5\}, \qquad (3.3.11)$$

$$x_{Topt} = -2e^2\{2[\pi_1+(\sigma_3+\sigma_4)h^2+(\tau_6+\tau_8+\tau_{11})h^4]F_1+[\pi_3+\pi_4+(\sigma_6+\sigma_8)h^2$$
$$+(\tau_{10}+\tau_{13})h^4]h^2+3(\sigma_1+\tau_3h^2+\tfrac{2}{3}\tau_4h^4)F_2-(\sigma_4+\tau_8h^2)h^2F_3$$
$$+4\tau_1F_4+\tau_4h^2.(2F_6-F_5\}, \qquad (3.3.12)$$

respectively. In fig. 3.~~2~~ 6 the functions F_1 to F_6 are shown for a unit radius aperture. Thus using equations (3.3.11,12) and fig. 3.~~2~~ 6 the variation of the optimum image planes for sagittal and tangential orientations of the spatial frequency components of an image may be determined.

v) Some Comparisons

We will now consider the expressions for optimum image planes according to the various criteria when the aberration expansion takes certain simple forms. Let a be the radius of the circular entrance pupil, α_o the radius of the paraxial exit pupil and α the radius of the (transformed) reference sphere pupil. Then recalling the transformation of coordinates involved in reducing the reference sphere pupil to a circle it follows that

$$\alpha = (1-x_o).\alpha_o \tag{3.3.13}$$

where $x_o = \frac{1}{2}e'^{-2}\alpha_o^{\ 2}(1+\frac{1}{4}e'^{-2}\alpha_o^{\ 2}+\frac{1}{8}e'^{-4}\alpha_o^{\ 4})+0(8)$, (3.3.14)

and $e' = e.(1+\nu)$. In §1.7 exact relations between the third order transverse aberration coefficients p_1 to p_5 and the fourth order retardation coefficients π_1 to π_5 were given. In these relations various powers of the product em^{*-1} occur. Now $em^{*-1} = (\alpha_o/a) = m_p$, where m_p is the paraxial magnification between pupils. This result is used in the following comparisons.

a) Spherical aberration

i) Suppose $\varepsilon = p_1\rho^3$ and $R = \pi_1\lambda^2$

Then $x_{Yopt} = x_{Zopt} = -\frac{2}{3}m^*p_1a^2$ (3.3.15a)

$$x_{Vopt} = -2e^2\pi_1\alpha^2$$

$$x_{Sopt} = x_{Topt} = -4e^2\pi_1F_1\alpha^2,$$

where F_1 is evaluated for unit radius aperture.

Since $\pi_1 = \frac{1}{4}m^*(em^{*-1})^{-2}e^{-2}p_1 = \frac{1}{4}(\frac{a^2}{\alpha o^2})m^*e^{-2}p_1$ it follows that

$$x_{Vopt} = -\tfrac{1}{2}m^*p_1(\alpha/\alpha o)^2.a^2 \tag{3.3.15b}$$

$$x_{Sopt} = -m^*p_1(\alpha/\alpha o)^2.F_1a^2. \tag{3.3.15c}$$

The relation derived here between x_{Yopt} and x_{Vopt} is a well known result. However the author knows of no cases where expressions for x_{Sopt} and x_{Topt} have been derived. It is therefore of value to examine these expressions in detail. Referring to fig. 3.2 it can be seen that $F_2 \rightarrow 2/3$ as $r \rightarrow 0$, has a minimum value of 4/9 and $\to 1$ as $r \to 1$. The corresponding image plane positions are

$$
\begin{aligned}
x_{Sopt} \to \quad & -\tfrac{2}{3}p_1(\alpha/\alpha o)^2.a^2 && \text{as } r \to 0 \\
& -\tfrac{1}{2}p_1(\alpha/\alpha o)^2.a^2 && \text{at } r \simeq 0.2, 0.55 \\
& -\ p_1(\alpha/\alpha o)^2.a^2 && \text{as } r \to 1
\end{aligned}
$$

Thus the position of the optimum image plane using the minimization of the difference function as a criterion moves from the geometric optimum of zero frequency to slightly less than the wavefront variance limit near $r = 0.4$ then back through the geometric optimum to a maximum shift which is twice that of the wavefront variance optimum. The behaviour of x_{Sopt} is completely described by the variation of F_1 in fig. 3.~~7~~6

ii) Consider now the case

$$\varepsilon = p_1\rho^3+s_1\rho^5+t_1\rho^7, \qquad \text{and} \quad R = \pi_1\lambda^2+\sigma_1\lambda^3+\tau_1\lambda^4.$$

Then

$$x_{Gopt} = -m^*(\tfrac{2}{3}p_1a^2+\tfrac{1}{2}s_1a^4+\tfrac{2}{5}t_1a^6), \tag{3.3.16a}$$

$$x_{Vopt} = -e^2(2\pi_1\alpha^2+\tfrac{9}{5}\sigma_1\alpha^4+\tfrac{8}{5}\tau_1\alpha^6), \tag{3.3.16b}$$

$$x_{Sopt} = -e^2(4\pi_1F_1\alpha^2+6\sigma_1F_2\alpha^4+8\tau_1F_4\alpha^6). \tag{3.3.16c}$$

It was shown in §1.7 that a simple rigorous relation between the coefficients s_j and σ_j $(j=1,2,\ldots 9)$ and between t_j and τ_j $(j=1,2,\ldots 14)$ did not exist. However numerical results presented in that chapter show that use of the approximate relations obtained from (1.7.4) give a good indication of the relations between the coefficients of spherical aberration. Furthermore if Sands W_o coordinates are used rather than $\bar{O}T$ coordinates then the approximate relations become exact relations when the constants involved are set to unity. Hence we have the following approximate expressions

$$x_{Vopt} \simeq -(\tfrac{1}{2}p_1\underset{\sim}{a}^2+\tfrac{3}{10}s_1\underset{\sim}{a}^4+\tfrac{1}{5}t_1\underset{\sim}{a}^6), \qquad (3.3.17a)$$

$$x_{Sopt} \simeq -(p_1F_1\underset{\sim}{a}^2+s_1F_2\underset{\sim}{a}^4+t_1F_4\underset{\sim}{a}^6), \qquad (3.3.17b)$$

where $\underset{\sim}{a} = (\alpha/\alpha o).a$. Comparison of (3.3.16a) with (3.3.17a) shows that in the presence of spherical aberration described by the first three orders of the aberration expansion,

$$x_{Gopt}-x_{Vopt} = -\tfrac{1}{5}(p_1a^2+s_1a^4+t_1a^6)\tfrac{1}{30}p_1a^2.$$

Consider now equation (3.3.17b). F_2 is 0.5 at r=0, reaches a minimum value of about 0.3 and then increases to 1, behaviour which is similar to that of F_1 with respect to the constants multiplying the fifth order aberration terms in (3.3.16a) and (3.3.17a). F_4 also shows similar behaviour. Thus from the curves of F_1, F_2 and F_4 we can see that x_{Sopt} lies between x_{Gopt} and x_{Vopt} over the range $0 \leqslant r \leqslant 0.65$, for cases when only spherical aberration of the first three orders is significant. It should be noted that in deriving this result the ratio $(\alpha/\alpha o)$ has been taken as unity. If the W_o coordinates are used this result will be strictly true.

Thus for the discussion of optimum image planes the variance of the aberration difference function is a quantity whose usefulness extends far beyond the range of spatial frequencies for which it provides a useful approximation to the m.t.f.

b) <u>Astigmatism and Field Curvature</u>

The presence of field curvature is indicated by non-planar surfaces of best-focus. When surfaces of best focus are defined for the tangential and sagittal orientations of frequency components of the image, and these surfaces do not coincide, then astigmatism is present and the distance between the focii is often used as a measure of astigmatism. Thus the expressions for optimum image planes provide a means of comparing the measures of astigmatism obtained from the various assessment quantities under this criterion.

Suppose that the transverse aberration is given by

$$\varepsilon_y = (3p_3+p_4)\rho V^2\cos\theta+(s_4+s_6\cos^2\theta)\rho^3V^2\cos\theta+s_{10}\rho V^4\cos\theta+(t_4+t_6\cos^2\theta)$$
$$\rho^5V^2\cos +(t_{11}+t_{12}\cos^2\theta)\rho^3V^4\cos\theta+t_{18}\ V^6\cos\theta,$$

$$\varepsilon_z = (p_3+p_4)\rho V^2\sin\theta+(s_5+s_6\cos^2\theta)\rho^3V^2\sin\theta+s_{11}\rho V^4\sin\theta+(t_5+t_6\cos^2\theta)\rho^5V^2\sin\theta$$
$$+(t_{13}+t_{14}\cos^2\theta)\rho^3V^4\sin\theta+t_{19}\rho V^6\sin\theta. \tag{3.3.18}$$

The corresponding wavefront retardation will then be

$$R = \pi_3\lambda\nu+\pi_4\mu^2+\sigma_3\lambda^2\nu+\sigma_4\lambda\mu^2+\sigma_6\lambda\nu^2+\sigma_8\mu^2\nu+\tau_3\lambda^3\nu+\tau_4\lambda^2\mu^2$$
$$+\tau_6\lambda^2\nu^2+\tau_8\lambda\mu^2\nu+\tau_{10}\lambda\nu^3+\tau_{11}\mu^4+\tau_{13}\mu^2\nu^2. \tag{3.3.19}$$

Then the optimum planes according to the various criteria are:-

$$x_{Yopt} = -m^*V^2\{3p_3+p_4+\tfrac{2}{3}(s_4+\tfrac{3}{4}s_6)a^2+s_{10}V^2+\tfrac{1}{2}(t_4+\tfrac{3}{4}t_6)a^4$$
$$+\tfrac{2}{3}(t_{11}+\tfrac{3}{4}t_{13})a^2V^2+t_{18}V^4\}, \tag{3.3.20a}$$

$$x_{Zopt} = -m^*V^2\{p_3+p_4+\tfrac{2}{3}(s_5+\tfrac{1}{4}s_6)a^2+s_{11}V^2+\tfrac{1}{2}(t_5+\tfrac{1}{4}t_6)a^4$$
$$+\tfrac{2}{3}(t_{12}+\tfrac{1}{4}t_{14})a^2V^2+t_{19}V^4\} \tag{3.3.20b}$$

$$x_{Gopt} = -m^*V^2\{2p_3+p_4+\tfrac{1}{3}(s_4+s_5+s_6)a^2+\tfrac{1}{2}(s_{10}+s_{11})V^2+\tfrac{1}{4}(t_4+t_5+t_6)a^4$$
$$+\tfrac{1}{3}(t_{11}+t_{12}+\tfrac{3}{4}t_{13}+\tfrac{1}{4}t_{14})a^2V^2+\tfrac{1}{2}(t_{18}+t_{19})V^4\}. \tag{3.3.20c}$$

$$x_{Vopt} = -h^2\{2\pi_3+\pi_4+(2\sigma_3+\sigma_4)\alpha^2+(2\sigma_6+\sigma_8)h^2+\tfrac{9}{5}(\tau_3+\tau_4)\alpha^4+(2\tau_6+\tau_8+\tfrac{3}{4}\tau_{11})\alpha^2h^2$$
$$+(2\tau_{10}+\tau_{13})h^4\}. \tag{3.3.21}$$

$$x_{Sopt} = -h^2\{2\pi_3+4\sigma_3F_1+2\sigma_4F_3+2\sigma_6h^2+6\tau_3F_2+4(\tau_4F_5+\tau_6F_1+\tfrac{1}{2}\tau_8F_3)h^2$$
$$+2\tau_{10}h^4\}, \qquad (3.3.22a)$$

$$x_{Topt} = -h^2\{2(\pi_3+\pi_4)+4[(\sigma_3+\sigma_4)F_1-\tfrac{1}{2}\sigma_4F_3]h^2+2(\sigma_6+\sigma_8)h^2+6\tau_3F_2+2\tau_4(2F_6-F_5)$$
$$+[4(\tau_6+\tau_8+\tau_{11})F_1-2\tau_8F_3-4\tau_{11}F_3]h^2+2(\tau_{10}+\tau_{13})h^4\}. \qquad (3.3.22b)$$

If we define $x_{Mopt} = \frac{1}{2}(x_{Sopt}+x_{Topt})$ then we have

$$x_{Mopt} = -h^2\{2(\pi_3+\pi_4)+2(2\sigma_3+\sigma_4)F_1-\sigma_4F_3+(2\sigma_6+\sigma_8)h^2+6\tau_3F_2+2\tau_4(F_5+F_6)$$
$$+4\tau_6F_1.h^2+2\tau_8F_1h^2+2\tau_{11}(F_1-F_3)h^2+(2\tau_{10}+\tau_{13})h^4\}. \qquad (3.3.22c)$$

The expressions for x_{Gopt}, x_{Vopt} and x_{Mopt} may be considered as defining the mean field according to the respective criteria. Using the equations of (1.7.2) we have

$p_3 \to \pi_4$ and $p_4 \to 2\pi_3-\pi_4$ so that $2p_3+p_4 \to 2\pi_3 + \pi_4$.

Hence if $\varepsilon_y = (3p_3+p_4)\rho V^2\cos\theta$ and $\varepsilon_z = (p_3+p_4)\rho V^2\sin\theta$ then the mean fields according to all three criteria are in agreement. Also

$$x_{Yopt} - x_{Zopt} = x_{Topt} - x_{Sopt}$$

so that the measures of astigmatism are also equal.

Using the approximate forms of (1.7.3) obtained by neglecting the lower order coefficients we have

$$s_4 \to 4\sigma_3+2\sigma_4, \qquad s_5 \to 4\sigma_3, \qquad s_6 \to 2\sigma_4$$

$$s_{10} \to 2(\sigma_6+\sigma_8), \qquad s_{11} \to 2\sigma_6. \qquad (3.3.23)$$

Then $\frac{1}{3}(s_4+s_5+s_6) \to \frac{2}{3}s_4 \to \frac{4}{3}(2\sigma_3+\sigma_4)$, and $\frac{1}{2}(s_{10}+s_{11}) \to 2\sigma_6+\sigma_8$. Hence if the expansion for $\underline{\varepsilon}$ consists only of the third and fifth order field-curvature terms, then the approximate relations above give

$$x_{Gopt} = -h^2\{2\pi_3+\pi_4+\frac{4}{3}(2\sigma_3+\sigma_4)a^2+(2\sigma_6+\sigma_8)h^2\}, \qquad (3.3.24)$$

Since the expressions for the variance of the wavefront aberration and of the difference function have been calculated assuming a circular pupil then the coefficients involved must be the transformed coefficients of 3.1. However inspection of the transformation equations shows that σ_3 and σ_4 are unaffected while σ_6 and σ_8 are modified by a term of order $(1/F^4)$, where F is the f-number of the system being considered. Remembering that we are only using approximate relations between the transverse aberration coefficients and the retardation coefficients for higher order terms, it is clear that no purpose is served by retaining these correction terms arising from the transformation of the exit pupil to a circle. However if the W_o-coordinates were used then the correction terms ought to be retained since, as mentioned earlier, these approximate relations then become exact. In our case, however, (3.3.23) and (3.3.24) are adequate.

Suppose that $\sigma_4=0$. Examination of the F_1 curve in fig. 3.2 then shows that x_{Mopt} approaches x_{Gopt} as $r\to 0$, and at $r\simeq 0.2$ and 0.55 $x_{Gopt} \simeq x_{Vopt}$. However if $\sigma_4\neq 0$ the situation is not so simple. The quantity $(2F_1-F_3)$ - see fig. 3.3 - multiplying $\sigma_4 a^2$ has the value 7/6 at $r=0$, 1.0 at $r\simeq 0.09$, 0.63 and reaches a minimum of 0.782 when $r\simeq 0.36$. The geometric limit is reached at $r\simeq 0.8$ and since F_1 also tends to the geometric limit at $r\simeq 0.8$, in the presence of third and fifth order field-curvature terms $x_{Mopt} = x_{Gopt}$ at $r\simeq 0.8$ and $x_{Mopt} \simeq x_{Vopt}$ at $r=0.6$. Apart from these two cases, the variation of x_{Sopt} depends on the relative values of the various field curvature coefficients, though for $0\leqslant r\leqslant 0.8$ the value of x_{Mopt} lies close to x_{Vopt} or x_{Gopt}.

A similar examination of the relation between mean planes in the presence of seventh order field curvature terms could be carried out, but the complication involved is considerable. It is, however, worth noting the various optimum planes when the aperture dependent terms in the expressions (3.3.20a), (3.3.21) and (3.3.22c) are zero. Then

$$x_{Gopt} = -h^2\{2\pi_3+\pi_4+(2\sigma_6+\sigma_8)h^2+(2\tau_{10}+\tau_{13})h^4\}$$

and $x_{Gopt} = x_{Mopt} = x_{Vopt}$.

Consider now the differences $x_{Yopt}-x_{Zopt}$ and $x_{Topt}-x_{Sopt}$.

From (3.3.20) and (3.3.22) we have

$$x_{Yopt}-x_{Zopt} = \Delta x_G = -m^*V^2\{2p_3+\tfrac{2}{3}(s_4-s_5+\tfrac{1}{2}s_6)a^2+(s_{10}-s_{11})V^2+\tfrac{1}{2}(t_4-t_5$$
$$+\tfrac{1}{2}t_6)a^4+\tfrac{2}{3}(t_{11}-t_{12}+\tfrac{3}{4}t_{13}-\tfrac{1}{4}t_{14})a^2V^2+(t_{18}-t_{19}(V^4\}$$

$$x_{Topt}-x_{Sopt} = \Delta x_M = -h^2\{2\pi_4+2\sigma_4(2F_1-F_2-F_3)+2\sigma_8h^2+2\tau_4(2F_6-3F_5)$$
$$+4\tau_4(F_6-\tfrac{3}{2}F_5)+4\tau_8(F_1-F_3)h^2+4\tau_{11}(F_1-F_3)h^2+2\tau_{13}h^2\}$$

(3.3.25)

Again using the correspondences derived from (1.7.3) and (1.7.4) we can write $\Delta x_G = -h^2\{2\pi_4+2\sigma_4a^2+2\sigma_8h^2+4\tau_4a^4+(\frac{14}{3}\tau_6-\frac{1}{3}\tau_8-\frac{8}{3}\tau_{11})a^2h^2+2\tau_{13}h^4\}$

(3.3.26)

Comparison of (3.3.25,26) shows agreement in the measure of astigmatism for vanishingly small aperture. However suppose only σ_4 is non-zero. Then the difference in astigmatism measures is determined by the quantity $1+F_2+F_3-2F_1$ where the functions F_1, F_2 and F_3 are shown in fig. 3.~~2~~6, and fig. 3.3 shows the form of $1+F_2+F_3-2F_1$.

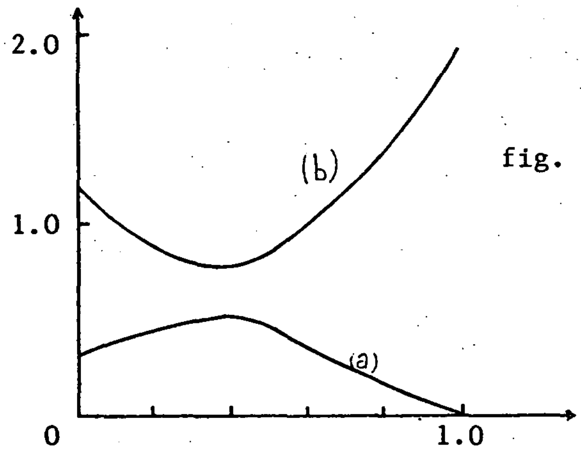


fig. 3.7 (a) $1 + F_2 + F_3 - 2F_1$

(b) $2F_1 - F_3$

Using the expressions of §2.5, 2.6 it is also possible to discuss transverse defocussing where, in (2.6.10), the term z_2 in the expression for b_{11} is the transverse defocussing coefficient.

§3.4 Orthogonal Expansion of the Transverse Aberration

In the previous section we have discussed the variation of the optimum image plane according to different criteria in the presence of certain forms of aberration. Now the choice of image plane is generally the final step in the design process; the image plane position is used as a final balancing parameter for those terms of the aberration expansion which vary with this position. The following four sections are devoted to a study of the earlier stage of design - namely the balancing of aberrations achieved by variation of the remaining constructional parameters of the system (curvatures, surface separations and glass choices).

Once the paraxial layout of a system is complete the designer normally reduces the third order aberrations to nominal values dictated by considerations of the proposed use of the system. The rest of the design process is one of reduction of types of image defect by the balancing of the various orders of aberration terms against one another. When the designer does not have access to a powerful automatic design program the determination of the form of balancing for a particular aperture and field angle poses considerable problems in the latter stages of fine correction of optical systems. The detailed analysis of spot diagrams and the introduction of the concept of the diapoint by Herzberger[9] give some idea of the nature of the general problem, while Cruickshank and Hills[10] have shown details of an actual aberration balancing procedure.

The expression of an assessment function as the sum of squares of coefficients in the aberration expansion provides an excellent means of determining the optimum form of balancing. The Zernike circle polynomials (orthogonal over the unit circle) were used by Nijboer[11] in expansions of the wavefront retardation function to provide an expression for the variance of the wavefront aberration. More recently Sumita[6] has used these polynomials to express the wavefront aberration difference function as a sum of squares of linear combinations of the classical aberration coefficients.

In this section the writer will show that $\overline{\varepsilon_y^{\,2}}$ and $\overline{\varepsilon_z^{\,2}}$ can be similarly treated and hence P_1 can be expressed in terms of squares of linear combinations of the classical aberration coefficients. P_2 may also be expressed in terms of the squares of these linear combinations, but since such an expression involves subtractions as well as additions the usefulness of the method is lost.

From (2.5.6) we have

$$\varepsilon_y = (A+\tfrac{1}{2}C+\tfrac{3}{8}E) + (B+\tfrac{3}{4}D)\cos\theta+\tfrac{1}{2}(C+E)\cos 2\theta+\tfrac{1}{4}D\cos 3\theta+\tfrac{1}{8}E\cos 4\theta,$$

$$\varepsilon_z = (\bar{B}+\tfrac{1}{2}\bar{D})\sin\theta+\tfrac{1}{4}(2\bar{C}+\bar{E})\sin 2\theta+\tfrac{1}{4}\bar{D}\sin 3\theta+\tfrac{1}{8}\bar{E}\sin 4\theta. \qquad (3.4.1)$$

Expressions for the coefficients $A,B,C,\ldots.\bar{E}$ are given in equations (2.5.7). Using the Zernike circle polynomials (3.4.1) can be written

$$\varepsilon_y = Y_{o}0+\frac{1}{3\sqrt{2}}\sum_{i=1}^{3} Y_{2i,o}R_{2i,o}(r) + \sum_{i=2}^{7}\sum_{j=1}^{k} Y_{ij}R_{ij}(r)\cos_j\theta,$$

$$\varepsilon_z = \frac{1}{\sqrt{2}}\sum_{i=1}^{3} Z_{2i,o}R_{2i,o}(r) + \sum_{i=2}^{7}\sum_{j=1}^{k} Z_{ij}R_{ij}(r)\sin\theta, \qquad (3.4.2)$$

where $(i-j)$ is even, k is the lesser of $(i, 8-i)$, and the radial polynomials, $R_{ij}(r)$, are given in table 3.1 below.

j \ i	0	1	2	3	4	5	6	7
0	1		$2r^2-1$		$6r^4-6r^2+1$		$20r^6-30r^4+12r^2-1$	
1		r		$3r^3-2r$		$10r^5-12r^3+3r$		$35r^7-60r^5+30r^3-4r$
2			r^2		$4r^4-3r^2$		$15r^6-20r^4+6r^2$	
3				r^3		$5r^5-4r^3$		$21r^7-30r^5+10r^3$
4					r^4		$6r^6-5r^4$	

Table 3.1

The radial polynomials are such that $0 \leq r \leq 1$ whereas the quantity ρ in equations (2.5.5) is the radial coordinate in focal length units. Hence the σ, μ and τ coefficients in (2.5.5) must be multiplied by a^s where $a = 1/2F$ and the values of s are given in table 3.2. The coefficients so obtained will be referred to as the normalized Buchdahl ray coefficients.

Coefficient \ s	0	1	2	3	4	5	6	7
p (σ)	5	4,3	2	1				
s (μ)	12	11,10	9,8,7	6,5,4	3,2	1		
t (τ)	20	19,18	17,16,15	14,13,12,11	10,9,8,7	6,5,4	3,2	1

Table 3.2

Expanding the radial polynomials in (3.4.2) according to table 3.1 and comparing the resultant expressions with (2.5.6,7) allows the coefficients Y_{ij}, Z_{ij} to be expressed in terms of the p, s and t coefficients and the field angle. These relations are given below:-

$$Y_{00} = (p_2+\tfrac{1}{3}s_2+\tfrac{1}{4}t_2)V+(p_5+\tfrac{1}{2}s_7+\tfrac{1}{3}t_7)V^3+(s_{12}+\tfrac{1}{2}t_{15})V^5+t_{20}V^7$$

$$\sqrt{2}\quad Y_{20} = (2p_2+s_2+\tfrac{9}{10}t_2)V+(s_7+t_7)V^3+t_{15}V^5,$$

$$Y_{11} = (\tfrac{2}{3}p_1+\tfrac{1}{2}s_1+\tfrac{2}{5}t_1)\ +(3p_3+p_4+\tfrac{2}{3}s_4+\tfrac{1}{2}s_6+\tfrac{1}{2}t_4+\tfrac{3}{8}t_6)V^2+(s_{10}+\tfrac{2}{3}t_{11}+\tfrac{1}{2}t_{12})V^4 +t_{18}V^6,$$

$$\sqrt{2}\quad Y_{40} = (\tfrac{1}{3}s_2+\tfrac{1}{2}t_2)V+\tfrac{1}{3}s_7V^3,$$

$$Y_{31} = (\tfrac{1}{3}p_1+\tfrac{2}{5}s_1+\tfrac{2}{5}t_1)+(\tfrac{1}{3}s_4+\tfrac{1}{4}s_6+\tfrac{2}{5}t_4+\tfrac{3}{10}t_6)V^2+(\tfrac{1}{3}t_{11}+\tfrac{1}{4}t_{12})V^4,$$

$$Y_{22} = (p_2+\tfrac{3}{4}s_3+\tfrac{3}{5}t_3)V+(s_8+\tfrac{3}{4}t_8)V^3+t_{16}V^5,$$

$$\sqrt{2}\quad Y_{60} = \tfrac{1}{10}t_2V,$$

$$Y_{51} = \tfrac{1}{10}(s_1+\tfrac{12}{7}t_1)+\tfrac{1}{10}(t_4+\tfrac{3}{4}t_6)V^2,$$

$$Y_{42} = \tfrac{1}{4}(s_3+\tfrac{4}{3}t_3)V+\tfrac{1}{4}t_8V^3,$$

$$Y_{33} = \tfrac{1}{4}(s_6+\tfrac{4}{5}t_6)V^2+\tfrac{1}{4}t_{12}V^4,$$

$$Y_{71} = \tfrac{1}{35}t_1, \qquad Y_{62} = \tfrac{1}{15}t_3V, \qquad Y_{53} = \tfrac{1}{20}t_6V^2,$$

$$Y_{44} = t_{10}V^3. \tag{3.4.3}$$

$$Z_{11} = (\tfrac{2}{3}p_1+\tfrac{1}{2}s_1+\tfrac{2}{5}t_1)+(p_3+p_4+\tfrac{2}{5}s_5+\tfrac{1}{3}s_6+\tfrac{1}{2}t_5+\tfrac{1}{4}t_6)V^2+(s_{11}+\tfrac{2}{3}t_{12}+\tfrac{1}{3}t_{14})V^4+t_{19}V^6,$$

$$Z_{31} = (\tfrac{1}{3}p_1+\tfrac{2}{5}s_1+\tfrac{2}{5}t_1)+(\tfrac{1}{3}s_5+\tfrac{1}{6}s_6+\tfrac{2}{5}t_5+\tfrac{1}{5}t_6)V^2+\tfrac{1}{3}(t_{13}+\tfrac{1}{2}t_{14})V^4,$$

$$Z_{22} = (p_2+\tfrac{3}{4}s_3+\tfrac{3}{5}t_3)V+(s_9+\tfrac{3}{4}t_9)V^3+t_{17}V^5,$$

$$Z_{51} = \frac{1}{10}(s_1+\frac{12}{7}t_1)+\frac{1}{10}(t_5+\tfrac{1}{2}t_6)V^2,$$

$$Z_{42} = \tfrac{1}{4}(s_3+\frac{4}{3}t_3)V+\tfrac{1}{4}t_9V^3,$$

$$Z_{33} = \tfrac{1}{4}(s_6+\frac{4}{5}t_6)V^2 +\tfrac{1}{4}t_{14}V^4,$$

$$Z_{71} = \frac{1}{35}t_1, \qquad Z_{62} = \frac{1}{15}t_3V$$

$$Z_{53} = \frac{1}{20}t_6V^2, \qquad Z_{44} = t_{10}V^3$$

$$Z_{2i,o} = Y_{2i,o} \quad \text{for } i = 1,2,3. \tag{3.4.4}$$

If defocussing is to be considered then

$$Y_{11} := Y_{11} + x'v'_{ak}, \qquad Z_{11} := Z_{11} + x'v'_{ak}.$$

Now the radial polynomials satisfy the relation

$$\int_0^1 R_{ij}(\rho).R_{kj}(\rho)\rho d\rho = \delta_{ik}/(2n+2), \tag{3.4.5}$$

and we also have

$$\pi^{-1}\int_0^{2\pi} \cos k\theta \cos \ell\theta d\theta = \pi^{-1}\int_0^{2\pi} \sin k\theta \sin \ell\theta d\theta = \begin{matrix} 0 & \text{if } k\neq\ell \\ 1 & \text{if } k=\ell\neq 0 \\ 2 & \text{if } k=\ell=0 \end{matrix}, \tag{3.4.6}$$

Hence $$(\pi a^2)^{-1}\int_0^1\int_0^{2\pi} \varepsilon_y^2 \rho d\rho d\theta = \tfrac{1}{2}\sum_{n,m} Y_{nm}^2/(n+1), \ n\text{-}m\geq 0 \text{ and even}, \tag{3.4.7a}$$

$0\leqslant n\leqslant 7$, $m\leqslant 8-n$

$$(\pi a^2)^{-1}.\int_0^1\int_0^{2\pi} \varepsilon_z^2 \rho d\rho d\theta = \tfrac{1}{2}\sum_{n,m} Z_{nm}^2/(n+1) \tag{3.4.7b}$$

and $$P_1 = \overline{\varepsilon_y^2} + \overline{\varepsilon_z^2} = \tfrac{1}{2}\sum_{n,m} (Y_{nm}^2 + Z_{nm}^2)/(n+1). \tag{3.4.7c}$$

Thus we have expressed the radius of gyration of the spot diagram as a sum of squares of linear combinations of the normalized transverse aberration coefficients. The centroid of the spot diagram is at Y_{00} and so by summing for n>0 in equations (3.4.7) the aberrations are measured from the centroid of the spot diagram. The contributions of comatic and astigmatic types of aberration terms are clearly distinguished in equations (3.4.7) - comatic terms occur only in $Z_{2i,j}$ and $Y_{2i,j}$ while astigmatic terms occur only in $Z_{2i-1,j}$ and $Y_{2i-1,j}$. By writing

$$Y_{nm} := (2n+2)^{-\frac{1}{2}} Y_{nm}, \qquad Z_{nm} := (2n+2)^{-\frac{1}{2}} Z_{nm}$$

the relative values of the various orthogonal coefficients may be more easily ascertained.

§ 3.5 Orthogonalization of the Wavefront Aberration

Using methods similar to those of the preceeding section relations between the retardation coefficients and the coefficients of an expansion of the wavefront retardation function in terms of orthogonal polynomials can be obtained. For the fourth order coefficients such relations have been given by Born and Wolf. The complete set of relations for terms up to the eighth order have been obtained by the writer and are presented here. The retardation coefficients of Buchdahl are again normalized to a unit radius aperture and table 3.3 gives the appropriate powers of the exit pupil radius needed to normalize these coefficients.

Coeff. \ value of s	1	2	3	4	5	6	7	8
π	5	4,3	2	1				
σ	9	8,6	7,5	4,3	2	1		
τ	14	13,10	12,9	11,8,6	7,5	4,3	2	1

Table 3.3

Let $r^2 = (\rho/\alpha)^2 = (y^2+z^2)/\alpha^2$ where is the exit pupil radius. Then we write

$$R(y,z,h) = A_{00} + \frac{1}{\sqrt{2}}\sum_{n=1}^{\infty} A_{no}R_{no}(r) + \sum_{n=1}^{\infty}\sum_{m=0}^{n} A_{nm}R_{nm}(r)\cos m\theta, \qquad (3.5.1)$$

where $n-m \geqslant 0$ and is even. Then we find

$$A_{00} = \tfrac{1}{2}(\tfrac{2}{3}\pi_1+\tfrac{1}{2}\sigma_1+\tfrac{2}{5}\tau_1)+\tfrac{1}{2}(\pi_3+\tfrac{1}{4}\pi_5+\tfrac{2}{3}\sigma_3+\tfrac{1}{3}\sigma_4+\tfrac{1}{2}\tau_3+\tfrac{1}{4}\tau_4)h^2+(\sigma_6+\tfrac{1}{4}\sigma_8+\tfrac{2}{3}\tau_6+\tfrac{1}{3}\tau_8-\tfrac{1}{2}\tau_{11})h^4+$$
$$+(\tau_{10}+\tfrac{1}{4}\tau_{13})h^6$$

$$\sqrt{2}\quad A_{20} = (\pi_1+\tfrac{9}{10}\sigma_1+\tfrac{4}{5}\tau_1)+(\pi_3+\sigma_3+\tfrac{9}{10}\tau_3+\tfrac{1}{4}\pi_4+\tfrac{1}{2}\sigma_4+\tfrac{9}{20}\tau_4)h^2+(\sigma_6+\tau_6+\tfrac{1}{4}\sigma_8+\tfrac{1}{2}\tau_8)h^4$$
$$+(\tau_{10}+\tfrac{1}{4}\tau_{13})h^6,$$

$$A_{11} = (\tfrac{2}{3}\pi_2+\tfrac{1}{2}\sigma_2+\tfrac{2}{5}\tau_2)h+(\pi_5+\tfrac{2}{3}\sigma_5+\tfrac{1}{2}\tau_5+\tfrac{1}{2}\sigma_7+\tfrac{3}{8}\tau_7)h^3+(\sigma_9+\tfrac{2}{3}\tau_9+\tfrac{1}{2}\tau_{12})h^5+\tau_{14}h^7,$$

$$\sqrt{2}\quad A_{40} = (\tfrac{1}{3}\pi_1+\tfrac{1}{2}\sigma_1+\tfrac{4}{7}\tau_1)+(\tfrac{1}{3}\sigma_3+\tfrac{1}{2}\tau_3+\tfrac{1}{6}\sigma_4+\tfrac{1}{4}\tau_4)h^2+(\tfrac{1}{3}\tau_6+\tfrac{1}{6}\tau_8+\tfrac{1}{2}\tau_{11})h^4,$$

$$A_{31} = (\tfrac{1}{3}\pi_2+\tfrac{2}{5}\sigma_2+\tfrac{2}{5}\tau_2)h+(\tfrac{1}{3}\sigma_5+\tfrac{2}{5}\tau_5+\tfrac{1}{4}\sigma_7+\tfrac{3}{10}\tau_7)h^3+(\tfrac{1}{3}\tau_9+\tfrac{1}{4}\tau_{12})h^5,$$

$$A_{22} = (\tfrac{1}{2}\pi_4+\tfrac{3}{8}\sigma_4+\tfrac{3}{10}\tau_4)h^2+(\tfrac{1}{2}\sigma_8+\tfrac{3}{8}\tau_8+\tfrac{3}{8}\tau_{11})h^4+\tfrac{1}{2}\tau_{13}h^6,$$

$$\sqrt{2}\quad A_{60} = \frac{1}{10}(\sigma_1+2\tau_1)+\frac{1}{10}(\tau_3+\tfrac{1}{2}\tau_4)h^2,$$

$$A_{51} = \frac{1}{10}(\sigma_2+\frac{12}{7}\tau_2)h+\frac{1}{10}(\tau_5+\frac{3}{4}\tau_7)h^3,$$

$$A_{42} = \frac{1}{8}(\sigma_4+\frac{4}{3}\tau_4)h^2+\frac{1}{8}(\tau_8+\tau_{11})h^4,$$

$$A_{33} = \tfrac{1}{4}(\sigma_7+\tfrac{4}{5}\tau_7)h^3+\tfrac{1}{4}\tau_{12}h^5,$$

$$\sqrt{2}\quad A_{80} = \tau_1/35, \qquad A_{71} = \tau_2 h/35, \qquad A_{62} = \tau_4 h^2/30$$

$$A_{53} = \tau_7 h^3/20, \qquad A_{44} = \tau_{11}h^4/8. \qquad (3.5.2)$$

It should be noted that the coefficients used here are the normalized transformed coefficients. Then, using (3.4.5) and (3.4.6), we have

$$\overline{R^2} = \sum_{n,m} A_{nm}^2/(2n+2). \tag{3.5.3}$$

But $\bar{R} = A_{00}$ so

$$\overline{R^2}-\overline{R}^2 = \tfrac{1}{2}\sum_{n=1}^{\infty} \sum_{m=0}^{n} A_{nm}^2/(n+1), \quad n-m\geq 0 \text{ and even.} \tag{3.5.4}$$

§3.6 Orthogonal Expansion of the Aberration Difference Function

In §2.6 the variance of the aberration difference function was introduced as an image assessment quantity related to the modulation transfer function, M(r,), where

$$M(r,\psi) = T(r,\psi)/T_o(r,\psi). \tag{3.6.1}$$

$T(r,\psi)$ is the o.t.f. of an optical system at line frequency r and line orientation ψ to the sagittal plane, and T_o is the o.t.f. for a similar system having no aberrations. Hopkins has shown that

$$|M(r,\psi)| \geq 1-\frac{2\pi^2}{\lambda^2}V(r,\psi) \tag{3.6.2}$$

where $V(r,\psi) = \overline{\Delta^2(s,t)}-\overline{\Delta(s,t)}^2$, $s=r\sin\psi$, $t=r\cos\psi$

provided the right-hand side of (3.6.2) is positive. Equality holds for small values of r. It is clear from (3.6.2) that an expression for $V(r,\psi)$ as a sum of squares of linear combinations of aberration coefficients would provide a valuable design tool. By approximating the sheared pupil integration region involved in the calculation of $V(r,\psi)$ by an ellipse, Sumita has succeeded in deriving such an expression for arbitrary line orientations ψ. Numerical checks have established that this approximation of the integration region gives no more than a few percent error in the area of this region over the whole frequency range. Sumita has expressed his equations in matrix notation and the details of some of the matrices have been provided. However, since explicit expressions are needed in the following work, the writer has found it necessary to derive all expressions ab initio using only the approximation of the integration region given by Sumita. Expressions are obtained for $V(r,\psi)$ at $\psi=0, \pi/2$.

a) <u>Sagittal Case, $\psi=0$</u>

From equations (2.6.12) we have

$$(s,0) = 4s\{A_{11}+A_{31}\rho^2+A_{51}\rho^4+A_{71}\rho^6+A_{33}\rho^2\cos^2\theta+A_{53}\rho^4\cos^2\theta\}\rho\cos\theta$$

$$+4s\{B_{22}+B_{42}\rho^2+B_{62}\rho^4+B_{44}\rho^2\cos^2\theta\}\rho^2\sin\theta\cos\theta, \qquad (3.6.3)$$

where the A and B coefficients are given by (2.6.13,14).

The sheared pupil region, shown in fig. 3.4,

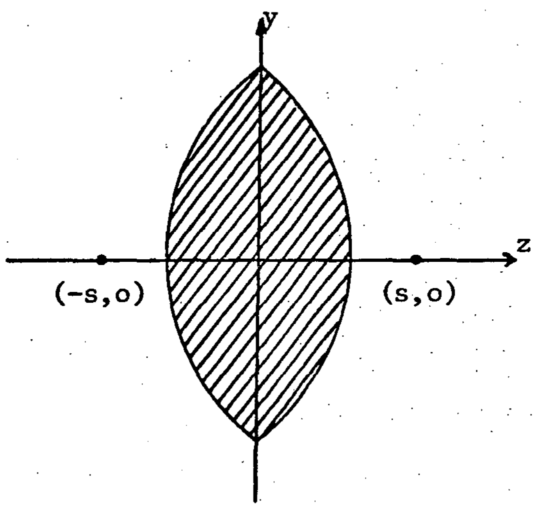


fig. 3.4

is transformed according to

$$\bar{y}: = y/\sigma_y, \quad \bar{z}: = z/\sigma_z \qquad (3.6.4)$$

where $(\bar{y},\bar{z})$ are the transformed coordinates of some point (y,z) in the region, and

$$\sigma_y = \pi^{-1}S/(a-r), \qquad \sigma_z = a-r. \qquad (3.6.5)$$

S is the area of the integration region and is given by

$$S = a^2(2\beta-\sin 2\beta), \qquad \text{where } \sigma = \arccos(r/a).$$

The result of the transformation (3.6.4) is to reduce the integration to a unit circle. Now let

$$\chi^2 = \sigma_y^{\ 2}-\sigma_z^{\ 2}, \qquad \bar{y} = \bar{\rho}\cos\phi, \qquad \bar{z} = \bar{\rho}\sin\phi.$$

Then substitution into (3.6.3) gives

$$\begin{aligned}\Delta(s,0) = {} & (\gamma_{11}\bar{\rho}+\gamma_{31}\bar{\rho}^3+\gamma_{51}\bar{\rho}^5+\gamma_{71}\bar{\rho}^7)\sin\phi+(\gamma_{33}\bar{\rho}^3+\gamma_{53}\bar{\rho}^5+\gamma_{73}\bar{\rho}^7)\sin\phi\cos^2\phi \\ & +(\gamma_{55}\bar{\rho}^5+\gamma_{75}\bar{\rho}^7)\sin\phi\cos^4\phi+\gamma_{77}\bar{\rho}^7\sin\phi\cos^6\phi+(\gamma_{22}\bar{\rho}^2+\gamma_{42}\bar{\rho}^4 \\ & +\gamma_{62}\bar{\rho}^6)\sin\phi\cos\phi+(\gamma_{44}\bar{\rho}^4+\gamma_{64}\bar{\rho}^6)\sin\phi\cos^3\phi+\gamma_{66}\bar{\rho}^6\sin\phi\cos^5\phi,\end{aligned} \tag{3.6.6}$$

where

$$\begin{aligned}
\gamma_{11} &= 4s\sigma_z A_{11}, & \gamma_{33} &= 4s\sigma_z(A_{31}\chi^2-A_{33}\sigma_y^{\ 2}),\\
\gamma_{31} &= 4s\sigma_z^{\ 3}(A_{31}+A_{33}), & \gamma_{53} &= 4s\sigma_z^{\ 3}[2A_{51}\chi^2+A_{53}(\chi^2-\sigma_z^2)],\\
\gamma_{51} &= 4s\sigma_z^{\ 5}(A_{51}+A_{53}), & \gamma_{55} &= 4s\sigma_z\chi^2(A_{51}\chi^2-A_{53}\sigma_z^{\ 2}),\\
\gamma_{71} &= 4s\sigma_z^{\ 7}A_{71}, & \gamma_{73} &= 12s\sigma_z^{\ 5}\chi^2A_{71}\\
\gamma_{77} &= 4s\sigma_z\chi^6A_{71}, & \gamma_{75} &= 12s\sigma_z^{\ 3}\chi^4A_{71}\\
\gamma_{22} &= 4s\sigma_y\sigma_z B_{22}, & \gamma_{44} &= 4s\sigma_y\sigma_z(\chi^2B_{42}-\sigma_z^{\ 2}B_{44}),\\
\gamma_{42} &= 4s\sigma_y\sigma_z^{\ 3}(B_{42}+B_{44}), & \gamma_{64} &= 4s\sigma_y\sigma_z^{\ 3}\chi^2.2B_{62},\\
\gamma_{62} &= 4s\sigma_y\sigma_z^{\ 5}B_{62}, & \gamma_{66} &= 4s\sigma_y\sigma_z\chi^4B_{62}.
\end{aligned} \tag{3.6.7}$$

Suppose we now expand $\Delta(s,0)$ in terms of Zernike circle polynomials. Then we have

$$\Delta(s,0) = b_{00} + \frac{1}{\sqrt{2}}\sum_{n=2}^{\infty} b_{no}R_{no}(\bar{\rho}) + \sum_{n=1\ m=0} b_{nm}R_{nm}(\bar{\rho})\sin m\phi, \tag{3.6.8}$$

Equating the right-hand sides of (3.6.6) and (3.6.8) allows the b_{ij} coefficients to be expressed in terms of the $\gamma_{k\ell}$ terms and hence, via (3.6.7), (2.6.13,14) and (2.6.10), in terms of the retardation coefficients. Thus, we have

$$
\begin{aligned}
b_{11} &= \gamma_{11}+\tfrac{2}{3}\gamma_{31}+\tfrac{3}{10}\gamma_{51}+\tfrac{7}{15}\gamma_{71}+\tfrac{1}{6}\gamma_{33}-\tfrac{1}{120}\gamma_{53}+\tfrac{53}{240}\gamma_{73}+\tfrac{23}{240}\gamma_{55}+\tfrac{1}{840}\gamma_{75}+\tfrac{7}{480}\gamma_{77},\\
b_{31} &= \tfrac{1}{3}(\gamma_{31}+\tfrac{1}{4}\gamma_{33}+\tfrac{6}{5}\gamma_{51}+\tfrac{1}{10}\gamma_{53}+\tfrac{1}{5}\gamma_{55}+\tfrac{7}{5}\gamma_{71}+\tfrac{51}{140}\gamma_{73}+\tfrac{3}{35}\gamma_{75}+\tfrac{7}{160}\gamma_{77}),\\
b_{33} &= \tfrac{1}{4}(\gamma_{33}+\tfrac{4}{5}\gamma_{55}+\tfrac{2}{3}\gamma_{73}+\tfrac{3}{14}\gamma_{75}+\tfrac{3}{16}\gamma_{77}),\\
b_{51} &= \tfrac{1}{10}(\gamma_{51}+\tfrac{1}{4}\gamma_{53}+\tfrac{1}{8}\gamma_{55}+2\gamma_{71}+\tfrac{1}{2}\gamma_{73}+\tfrac{1}{4}\gamma_{75}+\tfrac{5}{64}\gamma_{77}),\\
b_{53} &= \tfrac{1}{20}(\gamma_{53}+\tfrac{3}{4}\gamma_{55}+\tfrac{10}{7}\gamma_{73}+\tfrac{15}{14}\gamma_{75}+\tfrac{27}{28}\gamma_{77}),\\
b_{55} &= \tfrac{1}{16}(\gamma_{55}+\tfrac{6}{7}\gamma_{75}+\tfrac{15}{14}\gamma_{77}),\\
b_{71} &= \tfrac{1}{120}(4\gamma_{71}+\gamma_{73}+\tfrac{1}{2}\gamma_{75}+\tfrac{5}{64}\gamma_{77}),\\
b_{73} &= \tfrac{1}{84}(\gamma_{73}+\tfrac{3}{4}\gamma_{75}+\tfrac{9}{16}\gamma_{77}),\\
b_{75} &= \tfrac{1}{112}(\gamma_{75}+\tfrac{5}{4}\gamma_{77}),\\
b_{77} &= \tfrac{1}{64}\gamma_{77}.
\end{aligned}
\tag{3.6.9}
$$

$$
\begin{aligned}
b_{22} &= (\gamma_{22}+\tfrac{3}{4}\gamma_{42}+\tfrac{3}{5}\gamma_{62}+\tfrac{3}{8}\gamma_{44}+\tfrac{3}{10}\gamma_{64}+\tfrac{7}{80}\gamma_{60})/2,\\
b_{42} &= (\gamma_{42}+\tfrac{4}{3}\gamma_{62}+\tfrac{1}{2}\gamma_{44}+\tfrac{2}{3}\gamma_{64}+\tfrac{7}{36}\gamma_{66})/8,\\
b_{62} &= (\gamma_{62}+\tfrac{1}{2}\gamma_{64}+\tfrac{7}{48}\gamma_{66})/30,\\
b_{44} &= (\gamma_{44}+\tfrac{5}{6}\gamma_{64}+\tfrac{5}{9}\gamma_{66})/8,\\
b_{64} &= (\gamma_{64}+\tfrac{2}{3}\gamma_{66})/48,\\
b_{66} &= \gamma_{66}/32.
\end{aligned}
\tag{3.6.10}
$$

b) Tangential Case, $\psi = \pi/2$

From equations (2.6.16, 17) we have

$$\Delta(0,t) = 2t\{C_{00}+C_{20}\rho^2+C_{40}\rho^4+C_{60}\rho^6+C_{22}\rho^2\cos^2\theta+C_{42}\rho^4\cos^2\theta+C_{62}\rho^6\cos^2\theta$$
$$+C_{44}\rho^4\cos^4\theta\}+2t\{D_{11}+D_{31}\rho^2+D_{51}\rho^4+D_{71}\rho^6+D_{33}\rho^2\cos^2\theta+D_{53}\rho^4\cos^2\theta\}\cos\theta, \tag{3.6.11}$$

with the C and D coefficients given by (2.6.19,20). For this case we write

$$\sigma_y = a-t, \qquad \sigma_z = (\pi^{-1}S/(a-t), \qquad \chi^2 = \sigma_z^{\ 2}-\sigma_y^{\ 2}$$

and the transformation of the approximating ellipse to a circle then gives

$$(0,t) = \alpha_{00}+\alpha_{20}\bar{\rho}^{-2}+\alpha_{40}\bar{\rho}^{-4}+\alpha_{60}\bar{\rho}^{-6}+(\alpha_{22}\bar{\rho}^{-2}+\alpha_{42}\bar{\rho}^{-4}+\alpha_{62}\bar{\rho}^{-6})\cos^2\phi+(\alpha_{44}\bar{\rho}^{-4}+\alpha_{64}\bar{\rho}^{-6})\cos^4\phi$$
$$+\alpha_{66}\bar{\rho}^{-6}\cos^6\phi+(\alpha_{11}\bar{\rho}+\alpha_{31}\bar{\rho}^{-3}+\alpha_{51}\bar{\rho}^{-5}+\alpha_{71}\bar{\rho}^{-7})\sin\phi+(\alpha_{33}\bar{\rho}^{-3}+\alpha_{53}\bar{\rho}^{-5}+\alpha_{73}\bar{\rho}^{-7})$$
$$\sin\phi\cos^2+(\alpha_{55}\bar{\rho}^{-5}+\alpha_{75}\bar{\rho}^{-7})\sin\phi\cos^4\phi+\alpha_{77}\bar{\rho}^{-7}\sin\phi\cos^6\phi, \tag{3.6.12}$$

where

$$\alpha_{00} = 2tC_{00}, \qquad \alpha_{44} = 2t(C_{40}\chi^4-C_{42}\sigma_y^{\ 2}\chi^2+C_{44}\sigma_y^{\ 4}),$$

$$\alpha_{20} = 2t\sigma_y^{\ 2}(C_{20}+C_{22}), \qquad \alpha_{44} = 2t[3C_{60}\sigma_y^{\ 2}\chi^4+C_{62}\chi^2(\sigma_2^{\ 2}-3\sigma_y^{\ 2})],$$

$$\alpha_{40} = 2t\sigma_y^{\ 4}(C_{40}+C_{42}+C_{44}), \qquad \alpha_{66} = 2t\chi^4(C_{60}\chi^2-C_{62}\sigma_y^{\ 2}),$$

$$\alpha_{60} = 2t\sigma_y^{\ 6}(C_{60}+C_{62}),$$

$$\alpha_{22} = 2t(C_{20}\ ^2-C_{22}\ _y^{\ 2}),$$

$$\alpha_{42} = 2t(2C_{40}\chi^4-C_{42}\sigma_y^{\ 2}\chi^2+C_{44}\sigma_y^{\ 4}),$$

$$\alpha_{62} = 2t\sigma_y^{\ 4}[3C_{60}\chi^2+(2\sigma_z^{\ 2}-3\sigma_y^{\ 2})C_{62}], \tag{3.6.13}$$

$$\alpha_{11} = 2t\sigma_y D_{11}, \qquad \alpha_{33} = 2t\sigma_y(D_{31}\chi^2-D_{33}\sigma_y^{\ 2}),$$

$$\alpha_{31} = 2t\sigma_y^{\ 3}(D_{31}+D_{33}), \qquad \alpha_{53} = 2t\sigma_y^{\ 3}[2D_{51}\chi^2+D_{53}(\chi^2-\sigma_y^{\ 2})],$$

$$\alpha_{51} = 2t\sigma_y^{\ 5}(D_{51}+D_{53}), \qquad \alpha_{73} = 6t\sigma_y^{\ 5}\chi^2 D_{71},$$

$$\alpha_{71} = 2t\sigma_y{}^7 D_{71}, \qquad \alpha_{55} = 2t\sigma_y \chi^2 (D_{51}\chi^2 - D_{53}\sigma_y{}^2),$$

$$\alpha_{77} = 2t\sigma_y \chi^6 D_{71}, \qquad \alpha_{75} = 6t\sigma_y{}^3 \chi^4 D_{71}. \tag{3.6.14}$$

Now write

$$\Delta(0,t) = d_{00} + \sqrt{2}^{\,\frac{1}{2}} \sum_{n=0}^{\infty} d_{no} R_{no}(\bar{\rho}) + \sum_{n=1}^{\infty} \sum_{m=0}^{n} (d_{nm}\cos m\phi + e_{nm}\sin m\phi) . R_{nm}(\bar{\rho}), \tag{3.6.15}$$

Comparison of (3.6.12) and (3.6.15) shows that $d_{nm}=0$ for odd m, and $e_{nm}=0$ for even m. Equating like terms then gives expressions for the orthogonal coefficients in terms of the retardation coefficients:-

$$d_{00} = (2\alpha_{00}+\alpha_{20}+\tfrac{2}{3}\alpha_{40}+\alpha_{60}+\tfrac{1}{2}\alpha_{22}+\tfrac{2}{3}\alpha_{42}+\tfrac{1}{4}\alpha_{62}+\tfrac{1}{4}\alpha_{44}+\tfrac{3}{8}\alpha_{64}-\tfrac{7}{320}\alpha_{66}),$$

$$\sqrt{2}\; d_{20} = (\alpha_{20}+\alpha_{40}+\tfrac{9}{10}\alpha_{60}+\tfrac{1}{2}\alpha_{22}+\tfrac{1}{2}\alpha_{42}+\tfrac{9}{20}\alpha_{62}+\tfrac{3}{8}\alpha_{44}+\tfrac{27}{80}\alpha_{64}-\tfrac{3}{320}\alpha_{66}),$$

$$\sqrt{2}\; d_{40} = \tfrac{1}{3}(\alpha_{40}+\tfrac{3}{2}\alpha_{60}+\tfrac{1}{2}\alpha_{42}+\tfrac{3}{8}\alpha_{44}+\tfrac{3}{4}\alpha_{62}+\tfrac{9}{16}\alpha_{64}+\tfrac{3}{64}\alpha_{66}),$$

$$\sqrt{2}\; d_{60} = \tfrac{1}{10}(\alpha_{60}+\tfrac{1}{2}\alpha_{62}+\tfrac{3}{8}\alpha_{64}+\tfrac{1}{32}\alpha_{66}),$$

$$d_{22} = \tfrac{1}{2}(\alpha_{22}+\tfrac{3}{4}\alpha_{42}+\tfrac{3}{4}\alpha_{44}+\tfrac{3}{5}\alpha_{62}+\tfrac{3}{5}\alpha_{64}+\tfrac{9}{80}\alpha_{66}),$$

$$d_{42} = \tfrac{1}{8}(\alpha_{42}+\alpha_{44}+\tfrac{4}{3}\alpha_{62}+\tfrac{4}{3}\alpha_{64}+\tfrac{1}{4}\alpha_{66}),$$

$$d_{62} = \tfrac{1}{30}(\alpha_{62}+\alpha_{64}+\tfrac{3}{16}\alpha_{66}),$$

$$d_{44} = \tfrac{1}{8}(\alpha_{44}+\tfrac{5}{6}\alpha_{64}+\tfrac{5}{8}\alpha_{66}),$$

$$d_{64} = \tfrac{1}{48}(\alpha_{64}+\tfrac{3}{4}\alpha_{66}),$$

$$d_{66} = \tfrac{1}{32}\alpha_{66}, \tag{3.6.16}$$

$$e_{11} = \alpha_{11}+\frac{2}{3}\alpha_{31}+\frac{3}{10}\alpha_{51}+\frac{7}{15}\alpha_{71}+\frac{1}{6}\alpha_{33}-\frac{1}{120}\alpha_{53}+\frac{53}{240}\alpha_{73}+\frac{23}{240}\alpha_{55}+\frac{1}{840}\alpha_{75}+\frac{7}{480}\alpha_{77},$$

$$e_{31} = \frac{1}{3}(\alpha_{31}+\tfrac{1}{4}\alpha_{33}+\frac{6}{5}\alpha_{51}+\frac{1}{10}\alpha_{53}+\frac{1}{5}\alpha_{55}+\frac{7}{5}\alpha_{71}+\frac{51}{140}\alpha_{73}+\frac{3}{35}\alpha_{75}+\frac{7}{160}\alpha_{77}),$$

$$e_{33} = \tfrac{1}{4}(\alpha_{33}+\frac{4}{5}\alpha_{55}+\frac{2}{3}\alpha_{73}+\frac{3}{4}\alpha_{75}+\frac{3}{16}\alpha_{77}),$$

$$e_{51} = \frac{1}{10}(\alpha_{51}+\tfrac{1}{4}\alpha_{53}+\frac{1}{8}\alpha_{55}+2\alpha_{71}+\tfrac{1}{2}\alpha_{73}+\tfrac{1}{4}\alpha_{75}+\frac{5}{64}\alpha_{77}),$$

$$e_{53} = \frac{1}{20}(\alpha_{53}+\frac{3}{4}\alpha_{55}+\frac{10}{7}\alpha_{73}+\frac{15}{14}\alpha_{75}+\frac{27}{28}\alpha_{77}),$$

$$e_{55} = \frac{1}{16}(\alpha_{55}+\frac{6}{7}\alpha_{75}+\frac{15}{14}\alpha_{77}),$$

$$e_{71} = \frac{1}{120}(4\alpha_{71}+\alpha_{73}+\tfrac{1}{2}\alpha_{75}+\frac{5}{64}\alpha_{77}),$$

$$e_{73} = \frac{1}{84}(\alpha_{73}+\frac{3}{4}\alpha_{75}+\frac{9}{16}\alpha_{77}),$$

$$e_{75} = \frac{1}{112}(\alpha_{75}+\frac{5}{4}\alpha_{77}),$$

$$e_{77} = \frac{1}{64}\alpha_{77}. \qquad (3.6.17)$$

Thus the aberration difference function has been expressed in terms of polynomials orthogonal over the unit circle. The variance of the difference function can therefore be written.

$$V(r,0) = \tfrac{1}{2}\sum_{n=1}^{7}\sum_{m=0}^{k} b_{nm}^{2}, \quad \text{where } k = \text{lesser of } (n,8-n) \qquad (3.6.18)$$

$$v(r,\pi/2) = \tfrac{1}{2}\sum_{n=1}^{7}\sum_{m=0}^{k} (d_{nm}^{2}+e_{nm}^{2}) \qquad (3.6.19)$$

§3.7 Comparison of Forms of Balancing

The previous three sections have shown that a number of the merit functions introduced in chapter two of this work can be expressed as the sums of squares of linear combinations of the classical aberration coefficients. In design work these merit functions would be reduced to as small a value as possible. Hence each of these linear combinations must be reduced towards zero, and so an examination of these combinations reveal the relative values of certain aberration coefficients which must be approached if a well-corrected system is to be obtained. Rather than examine the form of balancing for each of the different functions, we will compare, as far as possible, the form of the orthogonal coefficients obtained from the different merit functions. The complexity of expressions for the difference function variance terms is such, however, that even for very simple forms of the aberration expansion, little value would be obtained from a comparative analysis using this quantity. Attention is therefore confined to the radius of gyration and the variance of the wavefront retardation.

Using the approximations arising from §1.7 - or considering the aberration expansion expressed in Sands' W_o coordinates - we can rewrite the expressions for the Y, Z coefficients given in (3.4.3,4) in terms of the retardation coefficients. These are:-

$$Y_{00} = (\pi_2+\sigma_2+\tau_2)h+(\pi_5+\sigma_5+\tfrac{3}{4}\sigma_7+\tau_5+\tfrac{3}{4}\tau_7)h^3+(\sigma_9+\tau_9+\tfrac{3}{4}\tau_{12})h^5+\tau_{14}h^7,$$

$$\sqrt{2}\ Y_{20} = (2\pi_2+3\sigma_2+\tfrac{18}{5}\tau_2)h+(2\sigma_5+\tfrac{3}{2}\sigma_7+3\tau_5+\tfrac{9}{4}\tau_7)h^3+(2\tau_9+\tfrac{3}{2}\tau_{12})h^5,$$

$$Y_{11} = \tfrac{8}{3}(\pi_1+\tfrac{9}{8}\sigma_1+\tfrac{6}{5}\tau_1)+(2\pi_3+2\pi_4+\tfrac{8}{3}\sigma_3+\tfrac{7}{3}\sigma_4+3\tau_3+\tfrac{5}{2}\tau_4)h^2+(2\sigma_6+2\sigma_8+\tfrac{8}{3}\tau_6+\tfrac{7}{3}\tau_8+2\tau_{11})h^4$$
$$+2(\tau_{10}+\tau_{13})h^6,$$

$$\sqrt{2}\ Y_{40} = (\sigma_2+2\tau_2)h+\tfrac{2}{3}(\sigma_5+\tfrac{3}{4}\sigma_7)h^3,$$

$$Y_{31} = \tfrac{4}{3}(\pi_1+\tfrac{9}{5}\sigma_1+\tfrac{12}{5}\tau_1)+\tfrac{4}{3}(\sigma_3+\tfrac{7}{8}\sigma_4+\tfrac{9}{5}\tau_3+\tfrac{3}{2}\tau_4)h^2+\tfrac{4}{3}(\tau_6+\tfrac{7}{8}\tau_8+\tfrac{3}{4}\tau_{11})h^4,$$

$$Y_{22} = (\pi_2+\tfrac{3}{2}\sigma_2+\tfrac{9}{5}\tau_2)h+(\sigma_5+\tfrac{3}{2}\sigma_7+\tfrac{3}{2}\tau_5+\tfrac{15}{8}\tau_7)h^3+(\tau_9+\tfrac{3}{2}\tau_{12})h^5,$$

$$\sqrt{2}\ Y_{60} = \tfrac{2}{5}\tau_2 h,$$

$$Y_{51} = \tfrac{3}{5}(\sigma_1+\tfrac{16}{7}\tau_1)+\tfrac{3}{5}(\tau_3+\tfrac{5}{6}\tau_4)h^2,$$

$$Y_{42} = \tfrac{1}{2}(\sigma_2+2\tau_2)h+\ (\tau_5+\tfrac{5}{4}\tau_7)h^3,$$

$$Y_{33} = \tfrac{1}{2}(\sigma_4+\tfrac{8}{5}\tau_4)h^2+\tfrac{1}{2}(\tau_8+2\tau_{11})h^4,$$

$$Y_{71} = \frac{8}{35}\tau_1, \qquad Y_{62} = \frac{1}{5}\tau_2 h, \qquad Y_{53} = \frac{1}{5}\tau_4 h^2, \qquad Y_{44} = \tfrac{1}{4}\tau_7 h^3. \tag{3.7.1}$$

$$Z_{11} = \frac{8}{3}(\pi_1+\frac{9}{5}\sigma_1+\frac{6}{5}\tau_1)+2(\pi_3+\frac{4}{3}\sigma_3+\frac{1}{3}\sigma_4+\frac{3}{2}\tau_3+\tfrac{1}{2}\tau_4)h^2+2(\sigma_6+\tau_8+\frac{4}{3}\tau_{11})h^4+2\tau_{10}h^6,$$

$$Z_{31} = \frac{4}{3}(\pi_1+\frac{9}{5}\sigma_1+\frac{12}{5}\tau_1)+\frac{4}{3}(\sigma_3+\tfrac{1}{4}\sigma_4+\frac{9}{5}\tau_3+\frac{3}{5}\tau_4)h^2+\frac{4}{3}(\tau_6+\tfrac{1}{4}\tau_8)h^4,$$

$$Z_{22} = (\pi_2+\frac{3}{2}\sigma_2+\frac{9}{5}\tau_2)h+(\sigma_5+\frac{3}{2}\tau_5+\frac{3}{8}\tau_7)h^3+\tau_9 h^5,$$

$$Z_{51} = \frac{3}{5}(\sigma_1+\frac{16}{7}\tau_1)+\frac{3}{5}(\tau_3+\frac{1}{3}\tau_4)h^2,$$

$$Z_{42} = \tfrac{1}{2}(\sigma_2+2\tau_2)h+\tfrac{1}{2}(\tau_5+\tfrac{1}{4}\tau_7)h^3,$$

$$Z_{33} = \tfrac{1}{2}(\sigma_4+\frac{8}{5}\tau_4)h^2+\tfrac{1}{2}\tau_8 h^4,$$

$$Z_{71} = \frac{8}{35}\tau_1, \qquad Z_{62} = \frac{1}{5}\tau_2 h, \qquad Z_{53} = \frac{1}{5}\tau_4 h^2, \qquad Z_{44} = \tfrac{1}{4}\tau_7 h^3. \tag{3.7.2}$$

Consider now three simple cases of the aberration function form:-

a) $R = \pi_1\lambda^2 + \sigma_1\lambda^3 + \tau_1\lambda^4.$

Then
$$Y_{11} = Z_{11} = \frac{8}{3}(\pi_1+\frac{9}{8}\sigma_1+\frac{6}{5}\tau_1),$$

$$Y_{31} = Z_{31} = \frac{4}{3}(\pi_1+\frac{9}{5}\sigma_1+\frac{12}{5}\tau_1),$$

$$Y_{51} = Z_{51} = \frac{3}{5}(\sigma_1+\frac{16}{7}\tau_1), \qquad Y_{71} = \frac{8}{35}\tau_1; \tag{3.7.3}$$

and
$$\sqrt{2}\,A_{20} = (\pi_1+\frac{9}{10}\sigma_1+\frac{4}{5}\tau_1),$$

$$\sqrt{2}\,A_{40} = \frac{1}{3}(\pi_1+\frac{3}{2}\sigma_1+\frac{12}{7}\tau_1),$$

$$\sqrt{2}\,A_{60} = \frac{1}{10}(\sigma_1+2\tau_1), \qquad \sqrt{2}\,A_{80} = \frac{1}{35}\tau_1. \tag{3.7.4}$$

One very interesting point emerges from a comparison of these expressions. It can be seen that the constants multiplying the coefficients in the Y_{ij} expressions here are larger than the corresponding constants in the A_{ij} expressions. Since the optimization process involves reducing each of the orthogonal coefficients towards zero this means that optimization of the radius of gyration is obtained by correction of the spherical aberration to zero at a smaller aperture than that implied by using the wavefront variance.

b) $R = \pi_2\lambda\mu+\sigma_2\lambda^2\mu+\sigma_5\lambda\mu\nu+\sigma_7\mu^3.$

This is the case of third and fifth order coma. The non-zero orthogonal coefficients are given by

$$Y_{00} = (\pi_2+\sigma_2)h+(\sigma_5+\tfrac{3}{4}\sigma_7)h^3,$$

$$\sqrt{2}\quad Y_{20} = 2(\pi_2+\tfrac{3}{2}\sigma_2)h+2(\sigma_5+\tfrac{3}{4}\sigma_7)h^3,$$

$$\sqrt{2}\quad Y_{40} = \sigma_2h+\tfrac{2}{3}(\sigma_5+\tfrac{3}{4}\sigma_7)h^3,$$

$$Y_{22} = (\pi_2+\tfrac{3}{2}\sigma_2)h+(\sigma_5+\tfrac{3}{2}\sigma_7)h^3, \qquad Y_{42} = \tfrac{1}{2}\sigma_2h; \tag{3.7.5}$$

$$A_{11} = \tfrac{2}{3}(\pi_2+\tfrac{3}{4}\sigma_2)h+\tfrac{2}{3}(\sigma_5+\tfrac{3}{4}\sigma_7)h^3,$$

$$A_{31} = \tfrac{1}{3}(\pi_2+\tfrac{6}{5}\sigma_2)h+\tfrac{1}{3}(\sigma_5+\tfrac{3}{4}\sigma_7)h^3,$$

$$A_{51} = \frac{1}{10}\sigma_2h, \qquad A_{33} = \tfrac{1}{4}\sigma_7h^3. \tag{3.7.6}$$

The quantity $\sigma_5+\frac{3}{4}\sigma_7$ is common to Y_{00}, Y_{20}, Y_{40}, A_{11} and A_{31}, so it is clear that, for larger field angles, it would be useful to have $\sigma_5\simeq-\frac{3}{4}\sigma_7$. It can also be seen that the coma correction for systems of small aperture but moderate field is not much different under either criterion. The differences will be somewhat more marked at small fields and apertures, if coma is significant at these field angles.

c) $R = \pi_3\lambda\nu+\pi_4\mu^2+\sigma_3\lambda^2\nu+\sigma_4\lambda\mu^2+\sigma_6\lambda\nu^2+\sigma_8\mu^2\nu.$

In this case we are considering astigmatic terms of third and fifth order. The non-zero orthogonal coefficients take the form

$$Y_{11} = 2(\pi_3+\pi_4+\tfrac{4}{3}\sigma_3+\tfrac{7}{6}\sigma_4)h^2+2(\sigma_6+\sigma_8)h^4,$$

$$Y_{31} = \tfrac{4}{3}(\sigma_3+\tfrac{7}{8}\sigma_4)h^2, \qquad Y_{33} = \tfrac{1}{2}\sigma_4 h^2$$

$$Z_{11} = 2(\pi_3+\tfrac{4}{3}\sigma_3+\tfrac{1}{3}\sigma_4)+2\sigma_6 h^4 = Y_{11}-2[(\pi_4+\tfrac{5}{6}\sigma_4)h^2+\sigma_8 h^4],$$

$$Z_{31} = \tfrac{4}{3}(\sigma_3+\tfrac{1}{3}\sigma_4)h^2 = Y_{31} - \tfrac{5}{6}\sigma_4 h^2, \qquad Z_{33} = \tfrac{1}{2}\sigma_4 h^2 = Y_{33}. \qquad (3.7.7)$$

$$A_{00} = \tfrac{1}{2}(\pi_3+\tfrac{2}{3}\sigma_3+\tfrac{1}{3}\sigma_4)h^2+(\sigma_6+\tfrac{1}{4}\sigma_8)h^4,$$

$$\sqrt{2}\ A_{20} = (\pi_3+\sigma_3+\tfrac{1}{2}\sigma_4)h^2+(\sigma_6+\tfrac{1}{4}\sigma_8)h^4,$$

$$\sqrt{2}\ A_{40} = \tfrac{1}{8}(\sigma_3+\tfrac{1}{2}\sigma_4)h^2$$

$$A_{22} = \tfrac{1}{2}(\pi_4+\tfrac{3}{4}\sigma_4)h^2+\tfrac{1}{2}\sigma_8 h^4, \qquad A_{42} = \tfrac{1}{8}\sigma_4 h^2. \qquad (3.7.8)$$

If none of the coefficients π_3, π_4, σ_3, σ_4, σ_6 or σ_8 are zero then it is clear that not all of the Y, Z coefficients or the A_{ij} coefficients can be zero simultaneously. It can be seen that both Y_{31} and Z_{31} cannot both be zero. However A_{40} could be zero and the ratio of σ_3 to σ_4 required for this to be so lies between the ratios necessary for Y_{31} and Z_{31}, respectively, to be zero. Suppose that we have A_{40} equal to zero, i.e. $\sigma_3 = -\sigma_4$. Then from the conditions that $Y_{11} = Z_{11} = 0$ and $A_{22} = A_{20} = 0$ we find that

$Y_{11} = Z_{11} = 0$ and $A_{22} = A_{20} = 0$ we find that

$$\pi_3+\pi_4+\tfrac{1}{2}\sigma_4+(\sigma_6+\sigma_8)h^2=0 \qquad (3.7.9a)$$

$$\pi_3-\tfrac{1}{3}\sigma_4+\sigma_6 h^2=0 \qquad (3.7.9b)$$

and $$\pi_4+\tfrac{3}{4}\sigma_4+\sigma_8 h^2=0 \qquad (3.7.10a)$$

$$\pi_3+(\sigma_6+\sigma_8)h^2=0 \tag{3.7.10b}$$

Taking equations (3.7.9a,b) we can write

$$\pi_4+\frac{5}{6}\sigma_4+\sigma_8h^2=0 \tag{3.7.9c}$$

and from equations (3.7.10a,b) we have

$$\pi_3+\pi_4+\frac{3}{4}\sigma_4+(\sigma_6+2\sigma_8)h^2=0 \tag{3.7.10c}$$

Equations (3.7.9c) and (3.7.10a) are very similar as also are (3.7.9a) and (3.7.10c). We can therefore see that although there is a difference in the form of aberration balancing in the presence of astigmatism accordong to the two criteria discussed here, when the correction state is good the differences are not very large.

This section of work has presented some simple examples of the differences and similarities in the form of aberration balancing according to two extreme criteria of image quality. The results of the previous sections have shown that the variance of the difference function is a very versatile merit function which, over the frequency range $0<r\lesssim 0.75$ covers both geometric and diffraction limits of correction. The complexity of the algebra has precluded a detailed analytic study of this quantity for aberration balancing. However the following section contains numerical comparisons being carried out on systems specified by their constructional parameters.

§3.8 Numerical Comparison of the Various Assessment Functions

A number of image assessment functions were introduced in chapter 2, and details necessary for their calculation using the first three non-zero orders of the aberration expansion were given. In the earlier sections of this chapter some of these assessment functions have been examined theoretically when the aberration function takes particular forms. However the o.t.f.-based assessment functions S_1, S_2 and S_3 could not be examined in this fashion since no closed form for the o.t.f. integral has been derived. Also, while the theoretical comparisons of the various assessment functions under certain simplified conditions give valuable insight into these functions, it is necessary to have some knowledge of how they compare when used on actual systems. In this section, therefore, the author presents data computed for a number of systems working at various field angles and apertures, and having a wide range of degrees of correction. From this collection of data a comparison of the various functions is presented and the value of spot diagram assessment and of assessment using the aberration difference function is highlighted.

Before presenting the data and discussion it is perhaps opportune to give some details regarding the computation of the various quantities. All the programs necessary were written in Algol for the Elliott 503 machine of the Hydro-University Computing Centre. The machine has 8K of main core, 16K of core backingstore and ∿250K of slow access disk files. All programming was done by the author, except for the computation of the $\bar{O}T$ coefficients; this used the programme of Ford, modified for file handling of data. Size limitations in the store necessitated splitting the computation into discrete programs, these being stored on disk together with intermediate data. During the computation entry to a particular program on disk was determined by initial key settings on the word generator, there being a limited amount of flexibility as to which types of computation were to be performed. Further key settings were used to control data input and some details of the computation (e.g. whether the distortion coefficients were set to zero, or whether the ray aberrations were referred to the centroid of the spot diagram). All computations were carried out for a circular entrance pupil and systems having unit focal length.

There are five main programs in the set used for the computation with an auxiliary program for plotting optimum image planes according to the different criteria. Storage limitations have necessitated setting upper bounds on the number of different values of certain variables which may be used. Hence on a given run of the assessment suite of programs there must be no more than 5 values of aperture, 5 field angles, 5 spatial frequencies and 20 image plane positions. A further limitation, dictated by considerations of computing time, is that no more than 2000 o.t.f. values are computed in a given run. The time required for a single o.t.f. value to be calculated is approximately 3 seconds when a 12 × 12 rectangular mesh is used. Obviously, therefore, the above limitation in the number of o.t.f. values is not serious.

We have given listings of a number of programs in the appendix. The naming of most programs was originally based on their use in a simple automatic design program described in the following chapter, and these names were retained when certain of these were modified for incorporation into the assessment suite. The functions of the various programs in this suite are briefly described below:-

Program One (P1A): A modification of Ford's program U260, this program reads in the specifications of the system to be examined and details of apertures, field angles, spatial frequencies and image plane positions at which assessments are to be made. The $\bar{O}T$ coefficients are calculated, printed out and stored.

Program Two (P5): This program, as listed, has expressions for calculating the quantities P_1, P_2 and P_3 given by the expressions of §2.5(c). The method of §2.5(a) for determining the pupil periphery is contained in the procedure "pupilit" and the procedure "fitellipse" gives a least-squares fit of an ellipse to the pupil periphery. In the actual assessment suite however, the entrance pupil is taken to be circular and is assumed to be coincident with the paraxial entrance pupil. The procedure "Effecab" gives the coefficients p_j, s_j and t_j of equation (1.7.1) in terms of the $\bar{O}T$ coefficients while procedure "best focus" uses Sands' expressions for the optimum image plane based on minimization of P_1.

Program Three (P8A): This program computes the W-coefficients from a knowledge of the $\bar{O}T$ coefficients and then uses these to calculate the retardation coefficients.

Program Four (VST): In this program the retardation coefficients, modified according to §3.2, are used to calculate the variance of the wavefront aberration and of the aberration difference function. Also the planes which minimize the variances are calculated for the apertures, field angles and spatial frequencies specified initially. An option is available which allows the image plane positions at which the assessment is to take place to be input at this stage, rather than use the set of positions originally specified.

Program Five (P12): This program computes the o.t.f. values at the apertures, field angles, spatial frequencies and image plane positions specified. A key option allows the exit pupil apertures, field angles and image plane positions to be respecified at this point - a useful facility when the time of computation of o.t.f. values is considered. Output includes all o.t.f. values computed in real and imaginary form and also in modulus and phase form. Also the functions S_1, S_2 and S_3 are calculated, assuming that the same image plane positions have been specified for both sagittal and tangential computations. The calculated o.t.f. values have been checked against the SIRA[13] values showing good agreement. Despite a number of requests for additional checks from established optical schools, no such data has been obtained.

The systems chosen for the comparison of assessment functions were taken mainly from Cox[14]. However a selection of additional systems was also examined, including the nine objectives used in the study of convergence of the retardation expansion presented in chapter one, and four astronomical telescope objectives specified by Narashimon[15], these latter systems having aspheric surfaces. Data are presented for an aplanatic Cassegrain objective designed by Narashimon for operation at a focal length of 1200 inches; for a triplet objective designed by R.E. Hopkins[16] and considered to be near optimum in performance and a comparable triplet designed by Cruickshank[17], and for nine systems selected from the catalogue given by Cox. Most systems were analysed at five different values of aperture, for five field angles at each aperture and using five spatial frequencies in all cases. The assessments are presented for a single image plane for each aperture, this plane being, in most cases, that which is approximately optimum in terms of the radius of gyration when each field angle has equal weighting. An exception is made for the aplanatic objective, where the image surface is generally considered to be curved. In most cases results are only presented for two or three apertures of a particular system and at three spatial frequencies, and o.t.f. - based functions have not been calculated for all field angles due to the cost of such computations. However the results presented are typical of those obtained and in themselves provide a substantial guide to the comparative value of the various assessment functions. Tables 3.4 to 3.15 provide the data for comparison of the assessment functions. For each system the functions measuring image spread have been tabulated together, and those measuring astigmatism have also been tabulated together. Thus in one group we have the radius of gyration, P_1, the variance of the wavefront aberration, Var, the o.t.f.-based function, S_1, and the function

$S_{1M} = (M(r,0).M(r,\pi/2))^{\frac{1}{2}}$ where $M(r,\psi)$ is found using the variance, $V(r,\psi)$, of the aberration difference function (see §2.6(b))); in another group we have the functions P_2, S_2 and $S_{2M}(=M(r,\pi/2)-M(r,0))$, again found using $V(r,\psi)$).

An examination of the tables shows that very often the rankings of the states of correction at different field angles for the same system are not dependent on the particular assessment function used. However the relative magnitudes of the assessment functions show little correlation. Some problems may arise from the use of the quantities S_1 and S_{1M} which involve the product of two signed quantities. When both of the multiplicands have negative signs the resultant value for S_1 or S_{1M} is not distinguishable from corresponding positive signed multiplicands. However the writer has found that by considering such quantities as negative, and (in the case of S_{1M}) adding a further negative quantity ≈ -1 to the value the modified values of S_{1M} show an excellent consistency of assessment with the spot radius of gyration when aberrations are fairly large. Work by many authors has shown the usefulness of the spot diagram in image assessment when the aberrations are large, and hence this modified quantity S_{1M} may be considered a valid criterion of image quality under those conditions. Since the m.t.f. is well approximated by $1-\frac{2\pi^2}{\lambda^2}V(r,\psi)$ for well-corrected systems S_{1M} must therefore be a useful image quality function under these conditions also. Hence the modified function S_{1M} provides us with an image quality criterion which correlates well with the traditional geometrical optics quantities for lower quality systems, and which, when the m.t.f. is greater than ≈ 0.7, gives an accurate and useful representation of the diffraction image of a system in terms of its spatial frequency components. Thus, even though the aberration difference function has obvious theoretical significance over only a fairly limited range of correction, the results presented here indicate that its value can be extended dramatically. These conclusions are fully supported by the theoretical investigations presented in earlier sections of this

System: Narashimon's Aplanatic Objective. $\lambda = 1.7917 \times 10^{-8} f'$.

Ap.	Angle	P_1	Var	$S_{1M}(r)$ r=0.201	0.402	0.603	0.804	0.930	$S_1(r)$ 0.201	0.402	0.603	0.804	0.930
0.0625 (F/8)	0°	5.99×10^{-9}	3.63×10^{-6}	0.751	0.505	0.283	0.101	0.020	1.0	1.0	1.0	1.0	1.0
	0.1°	1.73×10^{-7}	3.03×10^{-2}	0.510	0.200	0.115	0.071	0.019	0.639	0.206	0.230	0.676	0.948

Table 3.4

System: Cruickshank triplet $\lambda = 9.022 \times 10^{-7} f'$.

Ap.	Angle	$P_1 \times 10^5$	Var	$S_{1M}(r)$ r=0.027	0.046	0.055	S_1 0.027	0.046	0.055
F/6.7	0°	0.928	0.037	0.982	0.950	0.927	0.954	0.900	0.868
	4°	2.12	0.193	0.917	0.766	0.651	-	-	-
	7°	4.07	0.687	0.672	-0.452	-0.696	0.722	0.464	0.364
	10.5°	4.23	0.557	0.655	-0.316	-0.395	0.641	0.300	0.186
	14°	29.5	41.7	-2.31*	-13.7*	-20.6*	-	-	-
				r=0.021	0.034	0.041	0.021	0.034	0.041
F/5	0°	2.95	0.432	0.862	0.614	0.443	0.826	0.602	0.485
	4°	4.29	0.736	0.744	0.359	0.100	-	-	-
	7°	8.75	3.47	-0.735	-2.27*	-3.42*	0.548	0.282	0.200
	10.5°	10.2	3.75	-1.67*	-5.00*	-6.95*	0.280	0.039	-0.082
	14°	31.2	83.6	-8.77*	-25.7*	-37.0*	-	-	-

Table 3.5(a)

System: Cruickshank triplet. $\lambda = 9.022 \times 10^{-7} f'$.

Ap.	Angle	$P_2 \times 10^5$	$S_{2M}(r)$			$S_2(r)$		
			r=0.027	0.046	0.055	0.027	0.046	0.055
F/6.7	4°	1.96	0.137	0.365	0.515	-	-	-
	7°	3.35	0.403	1.02	1.40	0.275	0.413	0.421
	10.5°	-2.73	-0.236	-0.666	-0.956	-0.131	-0.148	-0.076
	14°	28.5	30.8	84.2	120	-	-	-
			r=0.021	0.034	0.041	0.021	0.034	0.055
F/5	4°	2.24	0.043	-0.034	-0.155	-	-	-
	7°	7.06	1.58	3.66	4.75	0.002	-0.253	-0.312
	10.5°	6.80	2.17	4.76	5.94	-0.093	0.050	0.174
	14°	29.4	29.2	82.4	119	-	-	-

Table 3.5(b)

System: R.E. Hopkins triplet. $\lambda = 9.022 \times 10^{-7} f'$.

Ap.	Angle	$P_1 \times 10^5$	Var	$S_{1M}(r)$			$S_1(r)$		
				r=0.041	0.068	0.082	0.041	0.068	0.082
F/10	0°	1.74	0.056	0.938	0.828	0.752	0.906	0.791	0.723
	4°	2.16	0.098	0.906	0.723	0.577	–	–	–
	7°	3.52	0.305	0.749	-0.248	-0.641	0.751	0.455	0.295
	10.5°	4.55	0.546	0.622	-0.215	-1.129*	0.630	0.288	0.163
	14°	8.60	1.16	0.344	-3.69*	-5.18*	–	–	–
				r=0.027	0.046	0.055	0.027	0.046	0.055
F/6.7	0°	1.59	0.047	0.955	0.882	0.835	0.922	0.829	0.777
	4°	2.57	0.355	0.874	0.659	0.508	–	–	–
	7°	5.59	1.75	0.362	-1.27*	-1.96*	0.525	0.213	0.118
	10.5°	8.20	3.56	-1.25*	-3.30*	-4.61*	0.260	0.133	0.113
	14°	11.6	3.99	-2.19*	-6.04*	-8.63*	–	–	–
				r=0.021	0.034	0.041	0.021	0.034	0.041
F/5	0°	2.20	0.111	0.909	0.745	0.632	0.857	0.660	0.552
	4°	3.81	1.48	0.711	0.234	-1.07*	–	–	–
	7°	9.20	7.71	-0.529	-3.65*	-5.24*	0.348	0.129	-0.038
	10.5°	14.6	17.4	-3.53*	-9.89*	-14.0*	0.150	0.132	0.074
	14°	14.4	8.30	-3.56*	-9.26*	-12.8*	–	–	–

Table 3.6(a)

System: R.E. Hopkins triplet. $\lambda = 9.022 \times 10^{-7} f'$.

Ap.	Angle	$P_2 \times 10^5$	$S_{2M}(r)$			$S_2(r)$		
			r=0.041	0.068	0.082	0.041	0.068	0.082
F/10	4°	2.00	0.155	0.418	0.592	-	-	-
	7°	2.95	0.347	0.932	1.32	0.246	0.494	0.581
	10.5°	-2.45	-0.173	-0.437	-0.600	-0.199	-0.271	-0.243
	14°	2.57	0.534	1.50	2.16	-	-	-
			r=0.027	0.046	0.055	0.027	0.046	0.055
F/6.7	4°	0.940	0.063	0.199	0.303	-	-	-
	7°	2.73	0.363	1.00	1.43	0.147	0.205	0.232
	10.5°	-3.49	-0.237	-0.727	-1.08	-0.195	-0.050	-0.007
	14°	-3.88	-0.684	-1.63	-2.16	-	-	-
			r=0.021	0.034	0.041	0.021	0.034	0.041
F/5	4°	0.394	0.046	0.068	0.057	-	-	-
	7°	6.34	1.75	4.35	5.91	0.049	0.135	0.265
	10.5°	9.98	5.11	12.5	16.8	-0.136	0.072	0.000
	14°	5.83	1.08	1.65	1.39	-	-	-

Table 3.6(b)

System: Cox 3.03. $\lambda = 2.3134 \times 10^{-5} f'$.

Ap.	Angle	P_1 10^5	Var	$S_{1M}(r)$			$S_1(r)$		
				r=0.068	0.137	0.274	0.068	0.137	0.274
F/6	0°	3.99	0.00124	0.996	0.985	0.951	0.915	0.818	0.626
	5°	3.74	0.00106	0.997	0.989	0.965	0.916	0.821	0.634
	10°	6.60	0.0155	0.992	0.974	0.932	0.911	0.808	0.611
	15°	9.20	0.214	0.983	0.948	0.865	0.902	0.785	0.567
	20°	11.5	1.26	0.975	0.918	0.757	0.883	0.727	0.454
				r=0.046	0.091	0.182	0.046	0.091	0.182
	0°	3.63	0.00083	0.998	0.992	0.975	0.946	0.881	0.752
	5°	4.70	0.00344	0.996	0.986	0.958	0.943	0.873	0.730
	10°	14.2	0.0544	0.962	0.875	0.646	0.910	0.779	0.540
	15°	25.3	0.672	0.796	-0.156	-0.921	0.802	0.630	0.403
	20°	29.8	4.29	-0.455	-0.644	-3.71*	0.685	0.492	0.217
				r=0.034	0.068	0.137	0.034	0.068	0.137
	0°	22.5	0.0677	0.898	0.644	-1.06*	0.865	0.632	0.284
	5°	17.8	0.0617	0.930	0.734	-0.444	0.896	0.707	0.359
	10°	28.7	0.275	0.823	0.422	-1.56*	0.813	0.588	0.255
	15°	64.5	3.17	-1.23	-1.37	-6.34*	0.612	0.433	0.180
	20°	99.4	21.8	-6.07*	-20.8*	-55.6*	0.421	0.230	-0.101

Table 3.7(a)

System: Cox 3.03. $\lambda = 2.313 \times 10^{-5}$ f´.

Ap.	Angle	$P_2 \times 10^5$	$S_{2M}(r)$			$S_2(r)$		
			r=0.068	0.137	0.274	0.068	0.137	0.274
	5°	3.35	0.0047	0.0168	0.0498	0.0044	0.0150	0.0350
	10°	4.01	0.0050	0.0154	0.0390	0.0042	0.0122	0.0242
	15°	-1.90	-0.0031	-0.0263	-0.124	-0.0073	-0.0345	-0.0999
	20°	10.8	0.0337	0.114	0.343	0.0564	0.158	0.286
			r=0.046	0.091	0.182	0.046	0.091	0.182
	5°	0.783	0.0002	0.0050	0.0377	0.0015	0.0098	0.0444
	10°	10.6	0.0380	0.112	0.264	0.0362	0.0882	0.132
	15°	23.0	0.319	0.922	1.97	0.215	0.278	0.165
	20°	22.3	1.01	2.82	5.51	0.127	0.132	0.319
			r=0.034	0.068	0.137	0.034	0.068	0.137
	5°	-14.6	-0.0895	-0.317	-0.978	-0.0851	-0.239	-0.376
	10°	17.2	0.101	0.188	-0.190	0.0628	-0.0721	-0.323
	15°	60.7	2.71	8.34	20.2	0.305	0.157	-0.156
	20°	79.2	13.4	41.5	102	0.106	0.169	0.282

Table 3.7(b)

System: Cox 3.05. $\lambda = 2.3134 \times 10^{-5} f'$.

Ap.	Angle	$P_1 \times 10^5$	Var	$S_{1M}(r)$ r=0.086	$S_{1M}(r)$ 0.342	$S_{1M}(r)$ 0.684	$S_1(r)$ 0.086	$S_1(r)$ 0.342	$S_1(r)$ 0.684
F/7.5	0°	0.642	1.41×10^{-5}	1.00	0.999	1.00	0.897	0.575	0.202
	5°	1.66	0.00184	0.995	0.995	0.966	0.897	0.572	0.200
	10°	5.23	0.138	0.994	0.955	0.954	0.892	0.549	0.191
	15°	6.55	2.16	0.983	0.882	0.870	0.878	0.498	0.148
	20°	9.14	18.8	0.943	0.694	0.733	0.804	0.255	-0.151
				r=0.068	0.274	0.547	0.068	0.274	0.547
F/6	0°	2.65	1.39×10^{-4}	0.999	0.996	0.993	0.918	0.655	0.336
	5°	2.80	0.00310	0.999	0.990	0.984	0.918	0.651	0.332
	10°	6.93	0.224	0.989	0.909	0.854	0.909	0.597	0.287
	15°	9.69	3.52	0.959	0.665	0.412	0.872	0.477	0.165
	20°	11.0	30.7	0.821	-0.272	-1.28*	0.691	-0.056	-0.137

Table 3.8(a)

System: Cox 3.05. $\lambda = 2.3134 \times 10^{-5} f'$.

Ap.	Angle	$P_2 \times 10^5$	$S_{2M}(r)$			$S_2(r)$		
			r=0.086	0.342	0.684	0.086	0.342	0.684
F/7.5	5°	0.971	0.0006	0.0072	0.0044	0.0001	0.0045	0.0012
	10°	3.45	0.0059	0.0445	0.0448	0.0050	0.0253	0.0108
	15°	4.93	0.0278	0.173	0.203	0.0315	0.114	0.0844
	20°	6.24	0.0299	0.0382	0.0048	0.103	0.285	0.303
			r=0.068	0.274	0.547	0.068	0.274	0.547
F/6	5°	-2.67	-0.0025	-0.0147	-0.0243	-0.0028	-0.0098	-0.0080
	10°	2.75	0.0052	0.0385	0.0608	0.0049	0.0254	0.0242
	15°	7.72	0.0674	0.457	0.713	0.0784	0.238	0.217
	20°	1.99	0.142	0.601	0.511	0.239	0.347	0.273

Table 3.8(b)

System: Cox 3.08. $\lambda = 2.3134 \times 10^{-5} f'$.

Ap.	Angle	$P_1 \times 10^4$	Var	$S_{1M}(r)$ r=0.086	0.171	0.342	$S_1(r)$ 0.086	0.171	0.342
F/7.5	0°	2.18	0.0173	0.902	0.676	0.186	0.810	0.553	0.213
	5°	1.67	0.0193	0.943	0.813	0.531	0.845	0.646	0.344
	10°	0.919	0.275	0.977	0.929	0.840	0.862	0.698	0.431
	15°	3.57	2.33	-0.292	-1.48	-2.26	0.453	-0.068	-0.124
	20°	8.63	10.32	-2.67	-2.93	-10.1*	0.098	0.055	-0.005
				r=0.068	0.137	0.274	0.068	0.137	0.274
F/6	0°	2.60	0.0392	0.854	0.493	-1.45*	0.788	0.466	0.061
	5°	2.02	0.0447	0.915	0.704	0.119	0.838	0.601	0.248
	10°	1.31	0.533	0.954	0.849	0.601	0.842	0.636	0.357
	15°	4.50	4.34	-0.840	-1.92	-1.94	0.392	-0.242	0.046
	20°	10.8	19.8	-3.18	-4.10*	-21.1*	0.191	-0.016	0.018
				r=0.046	0.091	0.182	0.046	0.091	0.182
F/4	0°	3.31	0.153	0.748	0.057	-3.19*	0.722	0.250	-0.135
	5°	2.74	0.221	0.841	0.413	-1.91*	0.781	0.449	-0.059
	10°	2.59	2.12	0.834	0.300	-0.649	0.663	0.334	-0.066
	15°	6.87	15.5	-1.51	-1.609	-9.69*	0.398	-0.206	0.086
	20°	16.1	73.4	-3.82	-13.0*	-53.8*	-0.114	0.031	-0.031

Table 3.9(a)

System: Cox 3.08. $\lambda = 2.3134 \times 10^{-5} f'$.

Ap.	Angle	$P_2 \times 10^4$	$S_{2M}(r)$			$S_2(r)$		
			r=0.086	0.171	0.342	0.086	0.171	0.342
F/7.5	5°	-0.539	-0.0146	-0.0498	-0.130	-0.009	-0.029	-0.048
	10°	0.536	0.0248	0.0738	0.156	0.054	0.130	0.195
	15°	3.39	1.04	3.44	8.71	0.599	0.650	0.394
	20°	8.47	9.93	33.1	86.1	0.633	0.227	0.011
			r=0.068	0.137	0.274	0.068	0.137	0.274
F/6	5°	-0.584	-0.0199	-0.0730	-0.224	-0.008	-0.036	-0.090
	10°	0.842	0.0522	0.164	0.388	0.118	0.268	0.341
	15°	4.23	1.686	5.81	16.5	0.638	0.676	0.239
	20°	10.6	15.8	55.0	160	0.491	0.117	0.059
			r=0.046	0.091	0.182	0.046	0.091	0.182
F/4	5°	0.250	-0.020	-0.090	-0.382	0.032	-0.022	-0.285
	10°	1.88	0.188	0.645	1.84	0.397	0.554	0.406
	15°	6.27	3.70	13.6	45.0	0.494	0.451	0.073
	20°	15.5	33.8	125	420	0.423	0.037	0.083

Table 3.9(b)

System: Cox 3.16. $\lambda = 2.3134 \quad 10^{-5} f'$.

Ap.	Angle	$P_1 \times 10^4$	Var	$S_{1M}(r)$			$S_1(r)$		
				r=0.086	0.171	0.342	0.086	0.171	0.342
F/7.5	0°	0.157	1.35×10^{-4}	0.999	0.997	0.992	0.897	0.786	0.571
	5°	0.167	3.68×10^{-5}	1.0	0.999	0.999	0.897	0.787	0.574
	10°	0.655	0.00219	0.991	0.973	0.934	0.889	0.765	0.535
	20°	1.12	1.20	0.958	0.879	0.725	0.845	0.665	0.383
	26°	2.49	24.2	-0.753	-1.57	-1.72	0.640	0.455	0.234
				r=0.046	0.091	0.182	0.046	0.091	0.182
	0°	2.17	0.0332	0.910	0.711	0.249	0.863	0.658	0.360
	5°	2.12	0.0403	0.907	0.686	0.102	0.862	0.645	0.310
	10°	2.10	0.0658	0.900	0.601	-0.724	0.853	0.604	0.213
	20°	6.85	9.90	-1.85	-2.67*	-10.6*	0.545	0.379	0.118
				r=0.034	0.068	0.137	0.034	0.068	0.137
	0°	4.94	0.388	0.485	-1.84*	-5.77*	0.552	0.127	0.114
	5°	4.17	0.345	0.590	-1.34*	-4.76*	0.631	0.162	-0.071
	10°	3.84	0.414	0.650	-0.370	-2.83*	0.689	0.302	0.168
	20°	22.5	67.3	-18.2*	-61.1*	-165*	0.322	0.182	0.006

Table 3.10(a)

System: Cox 3.16. $\lambda = 2.3134 \times 10^{-5} f'$.

Ap.	Angle	$P_2 \times 10^4$	$S_{2M}(r)$			$S_2(r)$		
			r=0.086	0.171	0.342	0.086	0.171	0.342
	5°	-0.116	-0.0003	-0.0007	-0.0007	-0.001	-0.001	0.000
	10°	-0.502	-0.0088	-0.0261	-0.0604	-0.009	-0.023	-0.036
	20°	-0.940	0.0050	-0.0371	-0.246	0.015	-0.012	-0.064
	26°	1.98	1.53	4.29	8.96	0.299	0.259	0.222
			r=0.046	0.091	0.182	0.046	0.091	0.182
	5°	-1.07	-0.0394	-0.126	-0.330	-0.041	-0.095	-0.107
	10°	-1.86	-0.319	-0.488	-1.47	-0.130	-0.337	-0.423
	20°	6.46	5.93	17.7	40.8	0.290	0.147	-0.074
			0.034	0.068	0.137	0.034	0.068	0.137
	5°	-2.34	-0.199	-0.691	-2.06	-0.155	-0.113	0.147
	10°	-1.83	-0.144	-0.833	-3.99	-0.164	-0.453	-0.084
	20°	20.7	58.3	187	487	0.123	0.054	-0.008

Table 3.10(b)

System: Cox 3.67. $\lambda = 2.3134 \times 10^{-5} f'$.

Ap.	Angle	$P_1 \times 10^4$	Var	$S_{1M}(r)$			$S_1(r)$		
				r=0.114	0.228	0.456	0.114	0.228	0.456
F/10	0°	1.35	0.00537	0.954	0.837	0.668	0.821	0.603	0.308
	5°	1.03	0.00338	0.972	0.898	0.791	0.837	0.646	0.356
	10°	0.600	0.00063	0.994	0.984	0.992	0.854	0.703	0.429
	15°	1.68	0.00524	0.940	0.866	0.785	0.796	0.601	0.326
	20°	3.05	0.0248	0.750	0.361	-1.165	0.625	0.313	0.073
				r=0.057	0.114	0.228	0.057	0.114	0.228
F/5	0°	6.91	0.440	0.031	-3.837*	-13.13*	0.198	-1.245*	0.054
	5°	6.03	0.372	0.202	-2.972*	-9.338*	0.352	-1.118*	-0.039
	10°	3.91	0.292	0.589	-1.459*	-3.811*	0.529	-0.049	-0.028
	15°	3.54	0.617	-0.580	-0.512	-3.291*	0.610	0.328	-0.089
	20°	6.45	2.23	-0.903	-6.943	-15.90*	0.425	0.269	0.144

Table 3.11(a)

System: Cox 3.67. $\lambda = 2.3134 \times 10^{-5} f'$.

Ap.	Angle	$P_2 \times 10^4$	$S_{2M}(r)$			$S_2(r)$		
			r=0.114	0.228	0.456	0.114	0.228	0.456
F/10	5°	-0.0246	-0.00065	-0.00209	-0.00586	-0.000	-0.001	-0.002
	10°	0.178	0.00234	0.00240	-0.0023	0.004	0.008	0.009
	15°	-0.545	0.0184	0.0290	0.0395	0.016	0.017	0.023
	20°	-1.59	0.0389	0.0292	0.00531	0.004	-0.052	-0.039
			r=0.057	0.114	0.228	0.057	0.114	0.228
F/5	5°	-1.53	-0.0808	-0.282	-0.903	-0.062	-0.145	0.135
	10°	-2.07	-0.0299	-0.250	-1.58	0.163	0.115	0.059
	15°	1.82	1.19	3.19	5.09	0.176	0.153	0.181
	20°	4.47	5.99	17.2	35.1	0.097	0.143	0.076

Table 3.11(b)

System: Cox 4.05. $\lambda = 2.3134 \times 10^{-5} f'$.

Ap.	Angle	$P_1 \times 10^4$	Var	$S_{1M}(r)$					$S_1(r)$				
				r=0.091	0.182	0.365	0.730		0.091	0.182	0.365	0.730	
F/8	0°	0.545	0.00119	0.993	0.976	0.937	0.968		0.884	0.755	0.514	0.156	
	5°	0.415	0.00610	0.996	0.989	0.974	0.989		0.884	0.757	0.520	0.138	
	10°	0.841	0.139	0.987	0.963	0.920	0.948		0.869	0.723	0.467	0.100	
	20°	2.86	5.31	0.839	0.433	-0.647	0.430		0.725	0.454	0.197	-0.048	
	25°	6.64	18.6	-1.05	-1.95	-2.18	-1.56		-	-	-	-	
				r=0.068	0.137	0.274	0.547	0.821	0.068	0.137	0.274	0.547	0.821
F/6	0°	0.487	0.00202	0.994	0.976	0.917	0.888	0.986	0.913	0.811	0.605	0.302	0.086
	5°	0.604	0.0202	0.992	0.976	0.942	0.943	0.988	0.903	0.786	0.569	0.230	-0.051
	10°	1.34	0.310	0.963	0.885	0.709	0.588	0.875	0.862	0.682	0.392	0.050	-0.058
	20°	4.29	8.82	0.583	-0.845	-0.942	-1.28*	-0.320	0.637	0.421	0.127	0.035	-0.009
				r=0.057	0.114	0.228	0.456	0.684	0.057	0.114	0.228	0.456	0.684
F/5	0°	0.298	0.00043	0.998	0.994	0.981	0.971	0.998	0.932	0.855	0.702	0.427	0.201
	5°	0.986	0.0494	0.978	0.933	0.848	0.805	0.848	0.894	0.748	0.492	0.108	0.146
	10°	2.02	0.561	0.907	0.684	-0.378	-0.791	-0.644	0.804	0.529	0.180	-0.075	-0.048
	20°	5.61	11.8	-0.268	-1.13	-2.98*	-5.76*	-2.89*	0.609	0.367	0.049	-0.017	-0.009
				r=0.046	0.091	0.182	0.365	0.547	0.046	0.091	0.182	0.365	0.547
F/4	0°	0.357	0.00067	0.997	0.989	0.972	0.970	0.983	0.945	0.878	0.748	0.529	0.331
	5°	1.51	0.164	0.940	0.799	0.462	0.116	0.136	0.841	0.597	0.234	-0.189	-0.139
	10°	2.77	1.26	0.800	0.243	-1.68*	-3.25*	-3.11*	0.669	0.276	-0.065	-0.077	-0.041
	20°	8.54	16.1	-1.18	-1.83	-5.69*	-13.8*	-13.4*	0.595	0.320	-0.004	-0.053	0.009

Table 3.12(a)

System: Cox 4.05. $\lambda = 2.3134 \times 10^{-5} f'$.

Ap.	Angle	$P_2 \times 10^4$	$S_{2M}(r)$					$S_2(r)$				
			r=0.091	0.182	0.365	0.730		0.091	0.182	0.365	0.730	
F/8	5°	0.319	0.0041	0.0129	0.0290	0.0150		0.0076	0.0209	0.0363	0.0323	
	10°	-0.651	-0.0147	-0.0459	-0.1064	-0.0807		-0.0059	-0.0147	-0.0140	0.0463	
	20°	2.21	0.2110	0.6246	1.293	0.6220		0.1697	0.2483	0.1705	0.0993	
	25°	6.01	2.118	6.174	12.15	3.735		-	-	-	-	
			r=0.068	0.137	0.274	0.547	0.821	0.068	0.137	0.274	0.547	0.821
F/6	5°	0.368	0.0069	0.0217	0.0534	0.0560	0.0084	0.022	0.0630	0.1260	0.1619	0.1126
	10°	-0.998	-0.0344	-0.1171	-0.3238	-0.5131	-0.1763	0.0005	-0.0006	0.0159	0.1422	0.1161
	20°	3.66	0.5465	1.712	4.140	5.136	0.9736	0.3028	0.2009	0.0855	0.0381	0.0189
			r=0.057	0.114	0.228	0.456	0.684	0.057	0.114	0.228	0.456	0.684
F/5	5°	0.0956	0.0057	0.0146	0.0209	-0.0230	-0.0575	0.0415	0.1136	0.2268	0.3226	0.2941
	10°	-1.52	-0.0782	-0.2809	-0.8589	-1.706	-1.290	0.0046	0.0003	0.0363	0.1748	0.1269
	20°	4.91	0.9470	3.020	7.686	12.00	7.792	0.2812	0.0807	-0.0648	-0.0334	0.0269
			r=0.046	0.091	0.182	0.365	0.547	0.046	0.091	0.182	0.365	0.547
F/4	5°	0.348	0.0234	0.0641	0.1113	0.0119	0.142	0.1171	0.2738	0.3886	0.4058	0.2791
	10°	-1.80	-0.0968	-0.3948	-1.443	-3.734	-4.298	0.1065	0.1448	0.1882	0.1565	0.0812
	20°	8.18	2.482	8.110	21.94	41.05	41.75	0.3095	0.0988	-0.0302	0.1063	0.0079

Table 3.12(b)

System: Cox 4.07. $\lambda = 2.3134 \times 10^{-5} f'$.

Ap.	Angle	$P_1 \times 10^4$	Var	$S_{1M}(r)$					$S_1(r)$				
				r=0.057	0.114	0.228	0.456	0.684	0.057	0.114	0.228	0.456	0.684
F/5	0°	1.60	0.00573	0.957	0.884	0.801	0.818	0.644	0.891	0.760	0.580	0.364	-1.14*
	5°	1.70	0.0399	0.953	0.865	0.734	0.704	0.555	0.877	0.715	0.469	0.155	-0.094
	10°	2.02	0.358	0.941	0.801	0.446	-0.401	-0.511	0.850	0.625	0.263	-0.139	-0.065
	15°	2.70	2.15	0.873	0.625	0.080	-1.70*	-1.65*	0.809	0.562	0.223	-0.048	-0.024
	20°	3.83	10.4	0.157	-1.14	-1.58	-1.85	-2.21	0.769	0.629	0.411	0.100	0.010
				r=0.046	0.091	0.182	0.365	0.547	0.046	0.091	0.182	0.365	0.547
F/4	0°	3.40	0.0928	0.752	0.160	-2.21*	-3.69*	-4.76*	0.729	0.383	0.180	0.079	0.011
	5°	3.27	0.151	0.776	0.095	-1.91*	-3.33*	-4.12*	0.747	0.398	0.184	0.095	0.018
	10°	3.20	0.618	0.777	-0.448	-1.40	-2.07	-2.16	0.761	0.425	0.170	-0.116	-0.071
	15°	4.97	2.90	-0.365	-1.558*	-4.34*	-7.58*	-9.66*	0.700	0.394	0.147	-0.063	-0.051
	20°	9.68	12.5	-1.66	-4.48*	-11.1*	-14.0*	-13.1*	0.600	0.420	0.175	-0.047	0.023

Table 3.13(a)

System: Cox 4.07. $\lambda = 2.3134 \times 10^{-5} f'$.

Ap.	Angle	$P_2 \times 10^4$	$S_{2M}(r)$					$S_2(r)$				
			r=0.057	0.114	0.228	0.456	0.684	0.057	0.114	0.228	0.456	0.684
F/5	5°	-0.682	-0.0267	-0.0731	-0.1360	-0.2252	-0.3897	-0.0028	0.0125	0.0914	0.2109	0.1887
	10°	-1.23	-0.0915	-0.2730	-0.5941	-0.9662	-1.181	-0.0347	-0.0314	0.1312	0.2833	0.1850
	15°	1.10	0.0339	0.0258	-0.0929	0.0186	0.4767	0.0145	-0.0126	0.0180	0.0966	0.0513
	20°	3.54	0.9253	2.366	4.261	5.922	7.016	0.1944	0.1581	0.0976	0.1544	0.0835
			r=0.046	0.091	0.182	0.365	0.547	0.046	0.091	0.182	0.365	0.547
F/4	5°	-1.47	-0.1300	-0.4334	-1.155	-2.191	-2.864	-0.0789	-0.1206	-0.0548	-0.0481	-0.0394
	10°	-2.22	-0.2900	-1.093	-3.290	-6.501	-7.740	-0.1889	-0.3786	-0.0915	0.2994	0.2477
	15°	3.27	0.8600	1.948	2.564	2.619	6.100	-0.0363	-0.2989	-0.1646	0.1299	0.1163
	20°	8.26	6.738	18.45	37.21	54.35	70.31	0.1736	0.1450	0.0691	0.0997	-0.0293

Table 3.13(b)

System: Cox 4.29. $\lambda = 2.3134 \times 10^{-5} f'$.

				$S_{1M}(r)$			$S_1(r)$		
Ap.	Angle	$P_1 \times 10^4$	Var	r=0.068	0.274	0.547	0.068	0.274	0.547
F/6	0°	0.621	0.00347	0.990	0.850	0.810	0.910	0.563	0.278
	5°	0.585	0.0169	0.994	0.943	0.949	0.907	0.585	0.259
	10°	1.29	0.279	0.972	0.804	0.738	0.872	0.443	0.097
	15°	1.96	1.39	0.937	0.241	-0.511	0.831	0.256	0.056
	20°	2.96	1.96	0.867	-0.125	-0.273	–	–	–
				r=0.057	0.228	0.456	0.057	0.228	0.456
F/5	0°	0.740	0.00638	0.986	0.782	0.567	0.920	0.568	0.273
	5°	0.848	0.0363	0.987	0.885	0.837	0.911	0.557	0.231
	10°	1.59	0.473	0.959	0.675	0.494	0.862	0.349	-0.121
	15°	2.22	2.10	0.925	-0.504	-1.05	0.830	0.212	0.022
	20°	4.72	2.61	0.528	-1.05	-2.21*	–	–	–
				r=0.034	0.137	0.274	0.034	0.137	0.274
F/3	0°	1.11	0.0116	0.975	0.639	0.394	0.931	0.532	0.304
	5°	1.84	0.241	0.945	0.283	-0.409	0.897	0.277	-0.234
	10°	3.85	1.74	0.703	-1.68*	-2.25*	0.782	-0.138	-0.212
	15°	9.74	5.05	-1.23	-11.5*	-20.2*	0.626	0.129	0.057

Table 3.14(a)

System: Cox 4.29. $\lambda = 2.3134 \times 10^{-5} f'$.

Ap.	Angle	$P_2 \times 10^4$	$S_{2M}(r)$			$S_2(r)$		
			r=0.068	0.274	0.547	0.068	0.274	0.547
F/6	5°	0.342	0.0028	0.0194	0.0182	0.0127	0.0790	0.1117
	10°	-0.584	-0.0096	-0.0690	-0.1498	0.0163	0.0983	0.1682
	15°	-1.72	-0.0907	-0.7412	-1.122	-0.0793	-0.2129	-0.0467
	20°	2.02	0.1254	0.4019	0.7050	-	-	-
			r=0.057	0.228	0.456	0.057	0.228	0.456
F/5	5°	0.437	0.0039	0.0405	0.0521	0.0238	0.1702	0.2403
	10°	-0.680	-0.0213	-0.1385	-0.3103	0.0277	0.1915	0.2765
	15°	-1.80	-0.1223	-1.165	-2.173	-0.1126	-0.2607	-0.0351
	20°	4.29	0.6437	3.639	5.723	-	-	-
			r=0.034	0.137	0.274	0.034	0.137	0.274
F/3	5°	1.19	-0.0506	-0.5573	-0.8261	0.0307	0.2659	0.4688
	10°	3.00	0.0912	-1.299	-3.873	0.0731	0.2887	0.4330
	15°	8.80	3.968	23.34	33.51	0.0873	-0.0507	0.0495

Table 3.14(b)

System: Cox 4.32. $\lambda = 2.3134 \times 10^{-5} f'$.

				$S_{1M}(r)$			$S_1(r)$		
Ap.	Angle	$P_1 \times 10^4$	Var	r=0.068	0.274	0.547	0.068	0.274	0.547
F/6	0°	0.712	0.00330	0.989	0.875	0.817	0.908	0.578	0.280
	4°	0.531	0.00205	0.994	0.931	0.902	0.912	0.609	0.298
	8°	0.837	0.00590	0.987	0.886	0.847	0.900	0.551	0.231
	10°	1.37	0.0167	0.965	0.696	0.584	0.872	0.422	0.122
	12°	2.06	0.0363	0.921	0.307	-0.085	0.818	0.266	-0.022
				r=0.057	0.228	0.456	0.057	0.228	0.456
F/5	0°	0.930	0.00761	0.981	0.778	0.553	0.916	0.562	0.266
	4°	0.721	0.00525	0.989	0.868	0.736	0.922	0.617	0.323
	8°	1.06	0.0117	0.979	0.796	0.640	0.903	0.526	0.219
	10°	1.68	0.0312	0.945	0.472	-0.051	0.859	0.347	0.080
	12°	2.56	0.0721	0.873	-1.21*	-2.11*	0.774	0.213	0.040
				r=0.041	0.163	0.326	0.041	0.163	0.326
F/3.6	0°	0.665	0.00034	0.997	0.990	0.990	0.950	0.787	0.586
	4°	0.773	0.00958	0.993	0.916	0.792	0.943	0.695	0.415
	8°	1.89	0.0749	0.936	-0.534	-1.20	0.862	0.156	0.046
	10°	2.83	0.173	0.851	-0.567	-2.68*	0.750	0.002	-0.098
	12°	4.10	0.397	0.681	-3.53*	-7.86*	0.583	-0.060	-0.126

Table 3.15(a)

System: Cox 4.32. $\lambda = 2.3134 \times 10^{-5} f'$.

Ap.	Angle	$P_2 \times 10^5$	$S_{2M}(r)$			$S_2(r)$		
			r=0.068	0.274	0.547	0.068	0.274	0.547
F/6	4°	4.10	0.0070	0.0717	0.103	0.007	0.052	0.045
	8°	2.11	0.0028	0.0336	0.0265	0.010	0.063	0.087
	10°	-5.95	-0.0094	-0.0973	-0.188	0.008	0.166	0.093
	12°	-7.38	-0.0039	-0.0980	-0.233	0.028	-0.017	0.080
			r=0.057	0.228	0.456	0.057	0.228	0.456
F/5	4°	5.37	0.0106	0.128	0.250	0.013	0.095	0.107
	8°	5.30	0.0133	0.138	0.207	0.028	0.153	0.181
	10°	-4.21	0.0071	0.0334	-0.032	0.040	0.076	0.088
	12°	6.62	0.0413	0.251	0.241	0.091	-0.076	-0.034
			r=0.041	0.163	0.326	0.041	0.163	0.326
F/3.6	4°	3.83	-0.0028	-0.0888	-0.260	0.003	-0.005	-0.006
	8°	-7.01	-0.0682	-1.125	-3.15	-0.011	-0.130	0.132
	10°	-9.41	-0.1118	-1.959	-5.70	0.016	-0.018	0.225
	12°	10.7	-0.0325	-1.632	-5.94	0.118	0.123	0.126

Table 3.15(b)

System: Cox 3.16 at aperture F/7.5. $x' \times 10^3 f'$. $\lambda = 2.3134 \times 10^{-5} f'$.

Angle	Freq.	x_G	x_{var}	x_o	x_M	x_{So}	x_{SM}	x_{To}	x_{TM}
0°	.086	-0.81	-0.61	-0.72	-0.71	-0.72	-0.71	-0.72	-0.71
	.171			-0.56	-0.63	-0.56	-0.63	-0.56	-0.63
	.342			-0.55	-0.56	-0.55	-0.56	-0.55	-0.56
	.684			-0.72	-0.71	-0.72	-0.71	-0.72	-0.71
5°	.086	-1.25	-1.11	-1.21	-1.19	-1.30	-1.27	-1.12	-1.10
	.171			-1.15	-1.13	-1.23	-1.20	-1.07	-1.05
	.342			-1.09	-1.06	-1.17	-1.13	-1.00	-0.99
	.684			-1.18	-1.19	-1.28	-1.28	-1.08	-1.09
10°	.086	-2.31	-2.28	-2.38	-2.30	-2.75	-2.65	-2.00	-1.94
	.171			-2.37	-2.28	-2.71	-2.61	-2.03	-1.95
	.342			-2.36	-2.27	-2.68	-2.58	-2.03	-1.96
	.684			-2.35	-2.29	-2.76	-2.70	-1.94	-1.88
20°	.086	-2.33	-1.85		-1.52		-3.62		0.59
	.171								
	.342				-2.06		-3.86		-0.25
	.684				-1.65		-3.93		0.63

Table 3.16

System: Cox 3.16 at aperture F/4. $x' \times 10^3 f'$. $\lambda = 2.3134 \times 10^{-5} f'$.

Angle	Freq.	x_G	x_{var}	x_o	x_M	x_{So}	x_{SM}	x_{To}	x_{TM}
0°	.046	-2.36	-1.84	-2.32	-2.20	-2.32	-2.20	-2.32	-2.20
	.091			-2.08	-2.08	-2.08	-2.08	-2.08	-2.08
	.182			-1.90	-1.88	-1.90	-1.88	-1.90	-1.88
5°	.046	-2.33	-2.01	-2.25	-2.26	-2.50	-2.50	-2.00	-2.02
	0.91			-2.26	-2.28	-2.45	-2.40	-1.93	-1.96
	.182			-2.05	-2.05	-2.25	-2.24	-1.85	-1.85
10°	.046	-1.93	-2.07	-1.98	-1.98	-3.15	-3.04	-0.80	-0.91
	.091			-2.10	-2.04	-3.15	-3.02	-1.05	-1.06
	.182			-2.17	-2.11	-3.07	-2.98	-1.27	-1.24
20°	.046	4.30	5.33		7.50		0.79		14.2
	.091				6.65		0.41		12.9
	.182				5.36		-0.18		10.9

Table 3.17

System: Cox 3.18 at aperture F/7.5. $x' \times 10^3 f'$. $\lambda = 2.3134 \times 10^{-5} f'$.

Angle	Freq.	x_G	x_{var}	x_o	x_M	x_{So}	x_{SM}	x_{To}	x_{TM}
0°	.086			-0.45	-0.46	-0.45	-0.46	-0.45	-0.46
	.342	-0.53	-0.40	-0.42	-0.37	-0.42	-0.37	-0.42	-0.37
	.684			-0.48	-0.47	-0.48	-0.47	-0.48	-0.47
5°	.086			-0.70	-0.71	-1.02	-1.03	-0.38	-0.38
	.342	-0.78	-0.64	-0.65	-0.61	-0.97	-0.93	-0.33	-0.28
	.684			-0.71	-0.70	-1.05	-1.02	-0.36	-0.38
10°	.086			-0.99	-0.95	-2.15	-2.06	-0.18	-0.16
	.342	-1.11	-0.89	-0.89	-0.85	-2.05	-1.97	-0.27	-0.27
	.684			-0.98	-0.95	-2.12	-2.05	-0.16	-0.16
15°	.086				0.04		-1.37		1.44
	.342	-0.31	0.09		0.13		-1.30		1.55
	.684				0.05		-1.35		1.44

Table 3.18

System: Cox 3.18 at aperture F/6. $x' \times 10^3 f'$. $\lambda = 2.3134 \times 10^{-5} f'$.

Angle	Freq.	x_G	x_{var}	x_o	x_M	x_{So}	x_{SM}	x_{To}	x_{TM}
	.068			-0.75	-0.72	-0.75	-0.72	-0.75	-0.72
0°	.274	-0.80	-0.61	-0.59	-0.57	-0.59	-0.57	-0.59	-0.57
	.821			-0.87	-0.87	-0.87	-0.87	-0.85	-0.87
	.068			-0.96	-0.95	-1.30	-1.27	-0.62	-0.63
5°	.274	-1.04	-0.84	-0.82	-0.81	-1.15	-1.13	-0.48	-0.48
	.821			-1.07	-1.10	-1.35	-1.41	-0.79	-0.79
	.068			-1.22	-1.18	-2.37	-2.28	-0.06	-0.08
10°	.274	-1.35	-1.07	-1.18	-1.12	-2.23	-2.15	0.12	0.08
	.821			-1.40	-1.32	-2.52	-2.38	-0.28	-0.25
	.068				-0.17		-1.54		1.20
15°	.274	-0.48	-0.07		-0.04		-1.44		1.36
	.821				-0.28		-1.58		1.03

Table 3.19

System: Cox 3.18 at aperture F/3. $x' \times 10^3 f'$. $\lambda = 2.3134 \times 10^{-5} f'$.

Angle	Freq.	x_G	x_{var}	x_o	x_M	x_{So}	x_{SM}	x_{To}	x_{TM}
	.034			-2.02	-1.99	-2.02	-1.99	-2.02	-1.99
0°	.068	-2.05	-1.75	-1.95	-1.95	-1.95	-1.95	-1.95	-1.95
	.137			-1.84	-1.85	-1.84	-1.85	-1.84	-1.85
	.034			-2.20	-2.17	-2.48	-2.46	-1.93	-1.89
5°	.068	-2.24	-1.93	-2.13	-2.13	-2.44	-2.42	-1.83	-1.85
	.137			-2.06	-2.04	-2.38	-2.33	-1.74	-1.74
	.034			-2.33	-2.27	-3.28	-3.19	-1.38	-1.34
10°	.068	-2.33	-2.05	-2.28	-2.24	-3.24	-3.18	-1.32	-1.29
	.137			-2.14	-2.15	-3.18	-3.12	-1.10	-1.17
	.034				-1.06		-1.97		-0.14
15°	.068	-0.87	-0.91		-1.05		-2.00		-0.09
	.137				-1.03		-2.01		0.04

Table 3.20

System: Cox 3.32 at aperture F/3.3. $x' \times 10^3 f'$. $\lambda = 2.3134 \times 10^{-5} f'$.

Angle	Freq.	x_G	x_{var}	x_o	x_M	x_{So}	x_{SM}	x_{To}	x_{TM}
	.076			-0.93	-0.95	-0.93	-0.95	-0.93	-0.95
0°	.152	-1.04	-0.83	-0.88	-0.88	-0.88	-0.88	-0.88	-0.88
	.304			-0.76	-0.78	-0.76	-0.78	-0.76	-0.78
	.076			-0.56	-0.56	-1.06	-1.02	-0.06	-0.10
5°	.152	-0.55	-0.53	-0.55	-0.55	-0.98	-0.98	-0.11	-0.12
	3.04			-0.52	-0.52	-0.94	-0.93	-0.10	-0.10
	.076			0.54	-0.56	-0.88	-0.84	1.95	1.96
10°	.152	0.76	0.33	0.35	0.37	-0.96	-0.92	1.66	1.66
	.304			0.19	0.18	-0.08	-1.02	1.45	1.39
	.076				1.83		0.84		2.82
15°	.152	2.20	1.05		1.25		0.53		1.96
	.304				0.61		0.13		1.09

Table 3.21

System: Cox 3.67 at aperture F/6.25. $x' \times 10^3 f'$. $\lambda = 2.3134 \times 10^{-5} f'$.

Angle	Freq.	x_G	x_{var}	x_o	x_M	x_{So}	x_{SM}	x_{To}	x_{TM}
0°	.071	-9.22	-7.46	-8.8	-8.54	-8.8	-8.54	-8.8	-8.54
	.143			-7.8	-7.92	-7.8	-7.92	-7.8	-7.92
	.570			-7.0	-7.51	-7.0	-7.51	-7.0	-7.51
5°	.071	-10.4	-8.71	-9.95	-9.83	-9.7	-9.65	-10.2	-10.0
	.143			-9.15	-9.20	-9.2	-9.03	- 9.3	- 9.36
	.570			-8.75	-8.70	-8.6	-8.56	- 8.9	- 8.93
10°	.071	-13.7	-12.3	-13.8	-13.6	-13.3	-12.8	-14.3	-14.4
	.143			-12.9	-12.9	-12.3	-12.2	-13.4	-13.5
	.570			-12.5	-12.3	-11.8	-11.5	-14.1	-13.0
15°	.071	-18.4	-17.7		-19.3		-17.4		-21.1
	.143				-18.4		-16.8		-19.9
	.570				-17.1		-15.9		-19.3
20°	.071	-23.0	-24.0		-26.0		-22.7		-29.4
	.143				-24.9		-22.0		-27.7
	.570				-23.8		-20.8		-26.7

Table 3.22

chapter.

It may be noted that little attention has been given in the preceeding discussion to the modified quantity S_1. In fact this modification has no real value, since if both sagittal and tangential o.t.f.'s have dropped below +0.2 there is virtually no significance in their actual values and so S_1 then has little value as an image assessment quantity.

The agreement between the various assessment quantities in their measurement of astigmatism is not as satisfactory as that between measurements of image spread. No clear trend is defined in relations between the functions, though quite often P_2 agrees with both S_{2M} and S_2 at the lower spatial frequencies. Also there is rarely any consistency between S_2 and the other functions if the value of the o.t.f.'s from which S_2 is determined are below about 0.3. One important point does emerge, however. That is that if the o.t.f. values are above about 0.4 and the astigmatism measure is >0.1 then the values of P_2 and S_2 agree on the sign of the astigmatism and, to a large extent, on the ranking of such states of correction. Hence, even though the details of astigmatism correction according to the different criteria may vary quite considerably, we can see that when moderately large differences in sagittal and tangential resolution exist these are indicated both in P_2 and S_2.

Having examined some of the general trends it is of interest to note some details arising from examination of the tabulated data. In table 3.7(a), at aperture F/6, it will be seen that all rankings are the same despite the small differences existing between assessments at angles 0°, 5° and 10°. A similar situation is found in table 3.8(a). However from table 3.9(a) it can be seen that the variance of the wavefront retardation is anomalous in its behaviour, a result found also in table 3.12 at aperture F/8. It might be expected that the wave aberration variance would agree with S_1 at higher frequencies while P_1 would agree with S_{1M} and S_1 at lower frequencies, in those cases where correction states are good but not well within the diffraction limit. This appears to be the case in table 3.11(a) at aperture 0.05, and table 3.14 at aperture 0.0833, indicating that the wave aberration variance may give a better indication of the resolution of high frequency components of an image than does the radius of gyration. However the author believes that the results here give no firm indication that this is so. Further examination of other data is needed before any clear indication of the emphasis placed by this variance on particular spatial frequency components can be obtained.

Finally we will examine the predicted location of the optimum image plane according to the various criteria. In §3.3(b) the field curvature in the presence of astigmatism was defined in terms of the image plane lying midway between the planes giving optimum performance for sagittal and tangential orientations of the image. We will now consider the variation of this plane according to the different criteria. Tables 3.16 to 3.22 present details of the optimum image plane position according to the various criteria. Frequencies are expressed as fractions of the cut-off frequency ($1/F\lambda$) and the image plane positions are expressed in units of $1/1000^{th}$ of the focal length. The optimum planes using the o.t.f. are determined using the modulus of the o.t.f. rather than the real part. Some inaccuracy arises in the determination since interpolation, with a fairly limited number of points, was used. However the results for the o.t.f. optima are still of quite useful accuracy. For the remaining criteria, optimum planes are found using the expressions given earlier in this chapter.

Examination of the tables reveals the following important points:-

1) The optimum planes x_G and x_V provide bounds within which x_M and x_0 tend to lie, though it must be emphasized that x_M and x_0 are not always confined within these bounds. In fact the variation of x_M and x_0 with spatial frequency is very close to that predicted in §3.4 when the simple cases of

spherical aberration alone, or astigmatism alone were discussed - that is, at low spatial frequencies x_M tends to the geometrical optics limit, x_G, while as the spatial frequency increases x_M tends to the physical optics limit x_V, passes beyond it slightly then reverses its direction of change and tends back beyond x_V and x_G.

2) The value of x_M is generally quite close to x_0 even when the m.t.f. is below the value, 0.7, for which the variance of the difference function provides a valid approximation to it. Table 3.16 provides the most significant case of deviation from this, seen in the 10° data, though in tables 3.17 and 3.20 it can be seen that x_M does not lie between x_G and x_V when both aperture and field-angle are moderately large. This in fact would also correspond to cases where the truncated aberration expansions are unreliable.

§3.9 Conclusion

In this chapter we have considered the practical use of the various assessment functions introduced earlier. Transformations which allow the retardation coefficients to be modified and the pupil to be treated as circular have been given. Expressions for the radius of gyration of the spot diagram as a sum of squares of orthogonal coefficients have been derived, these orthogonal coefficients being themselves linear combinations of the classical aberration coefficients of Buchdahl. Similar expressions have also been given for the variance of the wavefront aberration and, using an approximation due to Sumita, for the variance of the aberration difference function arising in o.t.f. theory. Theoretical comparisons of predicted optimum image planes have been presented, and a comparative examination of the orthogonal coefficients arising from various assessment functions has been given.

Finally, detailed numerical comparisons of the predictions of image quality using geometrical and physical optics criteria have been carried out. These have shown that

a) differences do exist between the predictions of image quality of a given system when different criteria are used,

b) the variance of the aberration difference function provides the designer with an extremely versatile image assessment function which, when modified as the writer suggests, can be used over virtually the entire range of correction states;

c) the functions P_2 and S_2 introduced for measuring astigmatism have a rather limited usefulness. A more satisfactory quantity (and one widely accepted) is the distance between sagittal and tangential focii. Expressions for this distance have been given both for spot diagram analysis, or difference function variance analysis.

References

3.1) Japan Optical Engineering Research Association, 7, 15 (1970).

3.2) J.B. DeVelis, J.O.S.A. 55, 165 (1965).

3.3) F. Zernike, Physica, 1, 689 (1934).

3.4) M. Born, and E. Wolf, Principles of Optics (Pergamon Press, Inc., New York, 1968), p. 467.

3.5) B. Tatian, J.O.S.A., 64, 1083 (1974).

3.6) H. Sumita, Jpn. J. Appl. Phys., 8, 1027 (1969).

3.7) P.J. Sands, J.O.S.A., 63, 582 (1973).

3.8) E.H. Linfoot, Physica, 24, 476 (1958).

3.9) M. Herzberger, Modern Geometrical Optics (Interscience Publishers, New York, 1958).

3.10) F.D. Cruickshank, and G.A. Hills, J.O.S.A., 50, 379 (1960).

3.11) B.R.A. Nijboer, Physica, 10, 679 (1943).

3.12) H.H. Hopkins, Proc. Phys. Soc. 70B, 449 (1957).

3.13) T.L. Williams, A. Ashton, M.J. Worral, S.I.R.A. Report, 1970.

3.14) A. Cox, A System of Optical Design (Focal Press, London, 1964).

3.15) M.L.N. Vyakaranam, M.Sc. Thesis, University of Tasmania, 1969.

3.16) R.E. Hopkins, Japan. J. Appl. Phys., 4, Suppl. 1, 60 (1965).

3.17) F.D. Cruickshank, private communication.

CHAPTER 4

AN EXAMINATION OF TWO AUTOMATIC DESIGN PROGRAMS

§4.1. Introduction

The first chapter of this thesis was concerned with the aberration theory of Buchdahl and, in particular, § 1.6 dealt with extensions of this theory to allow computation of the derivatives of the seventh order $\bar{O}T$ coefficients with respect to constructional parameters of the system. The general theory for obtaining the coefficient derivatives was developed by Buchdahl[1], but details associated with the effects of pupil shifts on coordinates had not been given. These details have, therefore, been derived and are presented in the appendices of this work. In chapter two a number of assessment functions were introduced and expressions for these given. Particular attention was given to the spot diagram quantities P_1, P_2 and P_3. Using an approximating ellipse to define the pupil periphery expressions for these functions were derived, and the programming for their calculation is contained in the listings of appendix 5. In this chapter use is made of these parts of the earlier work, together with expressions for the derivatives of P_1, P_2 and P_3 with respect to axial curvatures, to form a rather simple automatic design program. The successive stages of an optimization using this program are examined using the various assessment functions and spot diagrams are also presented for most stages. In addition a triplet design program due to Cruickshank[2] was used to produce a high quality design, and the stages of optimization using this program were examined in similar fashion to the first-mentioned optimization program. Cruickshank's program reduces selected aberration coefficients by varying different parameters used in the initial thin lens conditions, and, in the latter stages, by balancing of certain types of aberration terms against one another.

Originally it was intended that a comparison of the predictions of different types of assessment function be carried out using the derivatives of these assessment functions with respect to axial curvatures. By defining analogous functions, as was done in chapter two, and using these in similar programs the differences (if any) in assessments using these functions, and the effects of such assessments on the course of a design and its end point were to have been determined. The enormity of the task of writing such an automatic design program from scratch - particularly with the limited computer facilities then available - was not initially appreciated by the author. Considerable efforts were made to develop the necessary optical theory and the simple design program that has been written utilises the completed theory for optimization using spot diagram quantities. Attempts to develop expressions for the derivatives of the retardation coefficients have not been successful, numerical results indicating that the programmed expressions were approximately correct but certainly not exact. The development of such expressions would be a valuable development of the theory, and may possibly be best achieved using the approximations of § 3.1 and Sands' W_o coefficients[3].

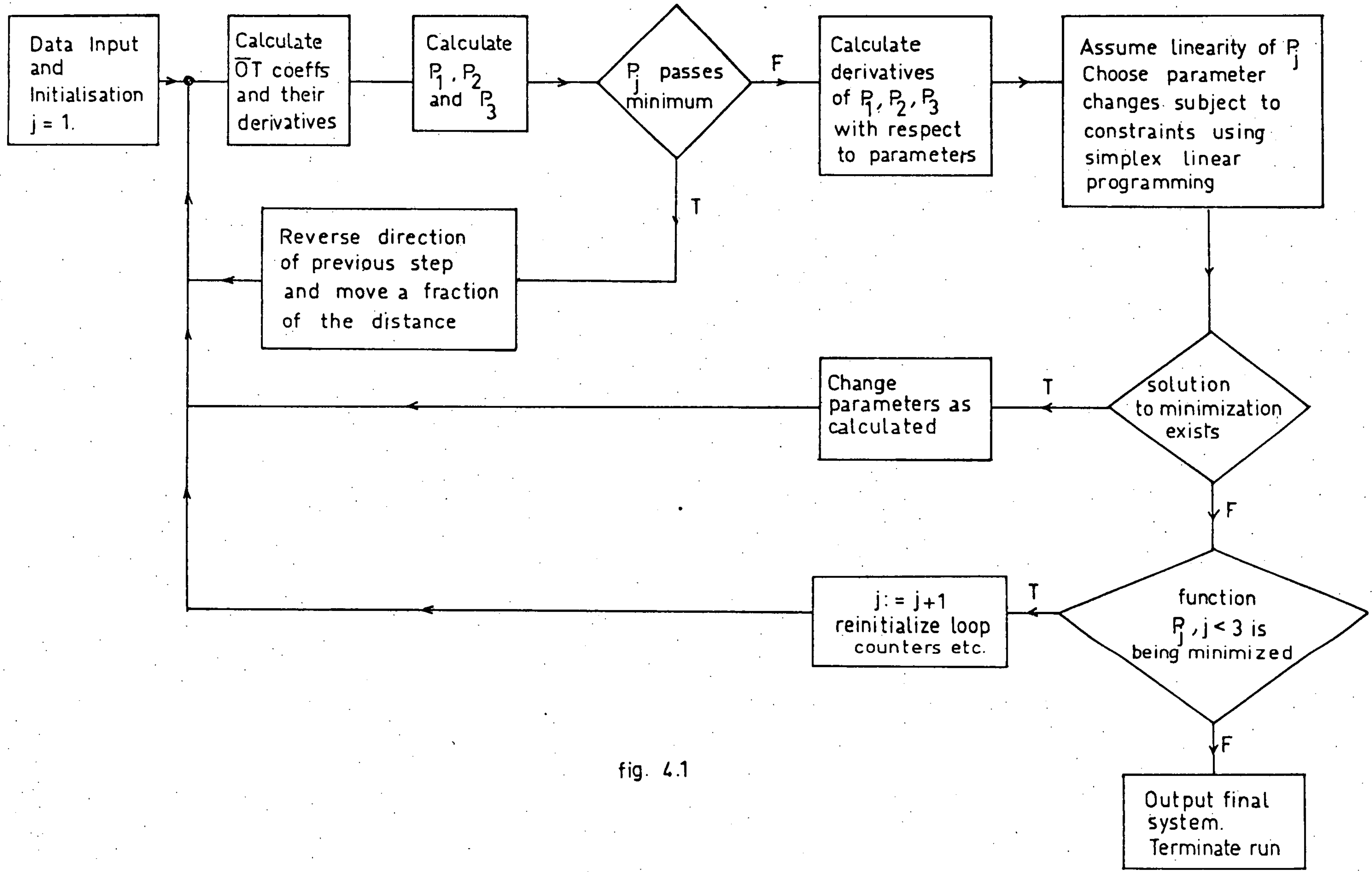
The problem of automatic design of optical systems using computers is one which has engaged optical designers and theoreticians, applied mathematicians, and computer programmers for quite some years, and an extensive literature on the subject has developed. Two key problems exist - firstly, to define a suitable merit function, and secondly, to devise highly efficient methods for minimizing this function with respect to its variables. The first task is essentially that of the optical designer and theorist, while the second is primarily the concern of the applied mathematician.

The merit function is some function of the constructional parameters of the system containing information on all aspects of the system with which one is concerned (e.g. image defects, glass thicknesses, overall length, sensitivity to tilt and decentering, cost, etc.). Most merit functions are expressed as the sum of positive quantities so that the automatic design process is reduced to that of minimizing a function of many variables. During the minimization process the function will normally be evaluated many times, and so it is advisable to choose as merit function some quantity which can be quickly calculated. For this reason quantities such as the radius of gyration of the spot diagram[4] or the wavefront variance[5,6] have been widely used. However King[7] has considered the use of the m.t.f., and quantities based on this have also been considered[8,9]. These quantities, of course, only assess the imaging properties of the system. The published literature contains little on sensitivity, though Grey[10] and Rimmer[11] have seriously tackled the problem. Details such as overall length and glass thickness often involve discussion of constraints which are incorporated into the merit function. In this work only that part of the merit function concerned with image defects has been considered.

Most methods of minimization of the assessment function appear to involve large fast machines with double precision packages readily available and using methods based predominantly on the use of functions calculated from raytrace data. The writer, however, has had access only to an Elliott 503 machine of 8K main store, 16K backing store and a 250K disk file unit. The word length is 39 bits and no double precision is available. Also the assessment functions used are calculated using the truncated aberration expansion rather than by raytrace. Given these conditions it was not feasible to use standard techniques to write an automatic design program, nor was it appropriate that such a task be attempted.

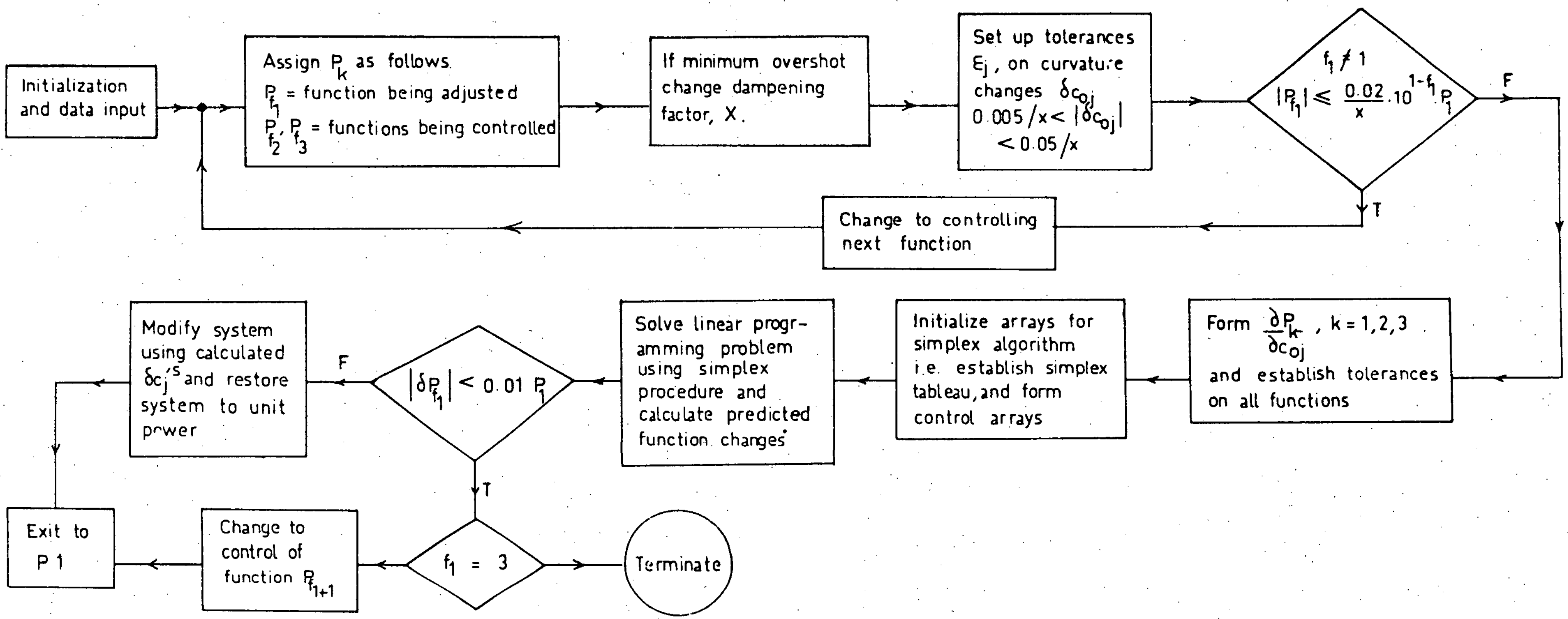
§4.2 A Simple Automatic Design Program Using Derivatives

The flow diagram of fig. 4.1 gives a general idea of the optimization program, while that of fig. 4.2 provides some detail on the decision making stage - i.e. the choice of parameter changes to continue the process of differential correction. The optimization proceeds by calculating the



Data Input and Initialisation j = 1.
Calculate $\overline{O}T$ coeffs and their derivatives
Calculate P_1, P_2 and P_3
P_j passes minimum
F
T
Calculate derivatives of P_1, P_2, P_3 with respect to parameters
Assume linearity of P_j Choose parameter changes subject to constraints using simplex linear programming
Reverse direction of previous step and move a fraction of the distance
Change parameters as calculated
T
solution to minimization exists
F
j: = j+1 reinitialize loop counters etc.
T
function P_j, j < 3 is being minimized
F
Output final system. Terminate run

fig. 4.1



Decision making in program PICK

fig. 4.2

aberration coefficients of the system and their derivatives with respect to axial curvatures. The pupil periphery is then determined using the method detailed in chapter two and an ellipse is fitted to this periphery using an iterative least squares method. The derivatives of the ellipse parameters are then found numerically. Following this the quantities P_j (j=1,2,3) are calculated. At this stage allowance is made for the possibility that the function, P_k, currently under adjustment has passed beyond the required value. If such an overshooting of the minimum has occurred then the sign of all previous parameter changes is reversed and the magnitude of the reverse step which is now taken is chosen to give a minimum assuming linearity of P_k. When overshooting does not occur the next step is to calculate the derivatives of the P_j quantities with respect to axial curvatures. Following this the appropriate curvature changes must be calculated.

The optimization strategy adopted is not conventional nor, probably, is it to be recommended as a general method. It was chosen to allow a comparison of assessment function influence on the course of automatic design, as mentioned in the introduction. Similarities to the strategy used by Cruickshank, and outlined in the following section, will be found. Firstly P_1 is reduced to near a minimum value, subject to the constraint that P_2 and P_3 do not increase significantly in magnitude. Then P_2 is reduced towards zero, while constraining P_1 and P_3; and finally P_3 is reduced as far as possible while constraining P_1 and P_2. The process may then be repeated, but the writer has found the small resultant improvement to be quite unjustifiable in terms of time taken. The process depends on the accuracy of the linearization of the functions P_j and this is controlled by restricting the magnitude of the curvature changes which are tolerated. The choice of curvature changes which will minimize the functions chosen, subject to constraints, was made using the linear programming simplex algorithm[12].

The programming of this automatic design process was carried out entirely by the author, except that the program used for computing the aberration coefficients was essentially a modification of an earlier program written by P.W. Ford. Because of limitations on program size the optimization process had to be written as a set of seven programs, these being stored as disk files and accessed in a pre-arranged sequence. Results and data from one program which were required by a later program were stored in data files and then re-accessed as required. The seven programs used approximately 65K of disk file space when dumped, this space not including array space. However it should be noted that a considerable amount of space was associated with file handling and reassigning values to variables which had to be redeclared in a number of programs. On a large machine considerable savings in machine space would be made. The functions of the various programs in the set are briefly outlined below, and the program name also given:-

program one (P1): Initializes certain control variables and inputs data. Then calculates the a and b $\bar{O}$T coefficients.

program two (P2): Uses Buchdahl's theory (L) to calculate the axial curvature derivatives of the paraxial, third and fifth order coefficients. Effects of the pupil shift are not incorporated into this part of the programming.

program three (P3): Calculate the derivatives of the seventh order
program four (P4): coefficients with respect to axial curvatures and then incorporates the effect of entrance pupil variation into the derivatives of all aberration coefficient derivatives calculated.

program five (P5): Determines the pupil periphery and a best-fitting ellipse to this periphery, then evaluates program five cont. numerically the derivatives with respect to axial curvatures of the ellipse parameters. Also the values of P_1, P_2 and P_3 are determined for some specified set of field angles and weighted sums of P_1, P_2 and P_3 across the field are formed.

program six (P6): Uses the derivatives of the aberration coefficients and pupil parameters to calculate derivatives of the weighted sums of P_1, P_2 and P_3.

program seven (PICK): Establishes the constraints on parameter changes required for approximate linearization, and also constraints on variation of the functions P_j ($j \neq k$). Then using the linear programming algorithm simplex, the parameter changes needed to minimize P_k subject to these various constraints are found. Program P_1 is then re-entered, omitting the initialization and data input section.

This automatic design program has been used to optimize the triplet designated Σ_1 used by Buchdahl for numerical illustrations in his monograph L and subsequent papers OACI-VIII. The triplet was to be optimized for an aperture F/4, and the field weighting was determined by evaluating the functions P_j ($j=1,2,3$) at angles of 0°, 5° and 7.5° and assigning equal weight to each angle. The process of optimization was examined by assessing the system resulting at the end of each cycle. After 10 cycles the program was ready to terminate or start again on optimizing P_1, and so this was taken to be sufficiently optimized. The spot diagrams in figs. 4.3 to 4.9 show visually the changes during the course of optimization.

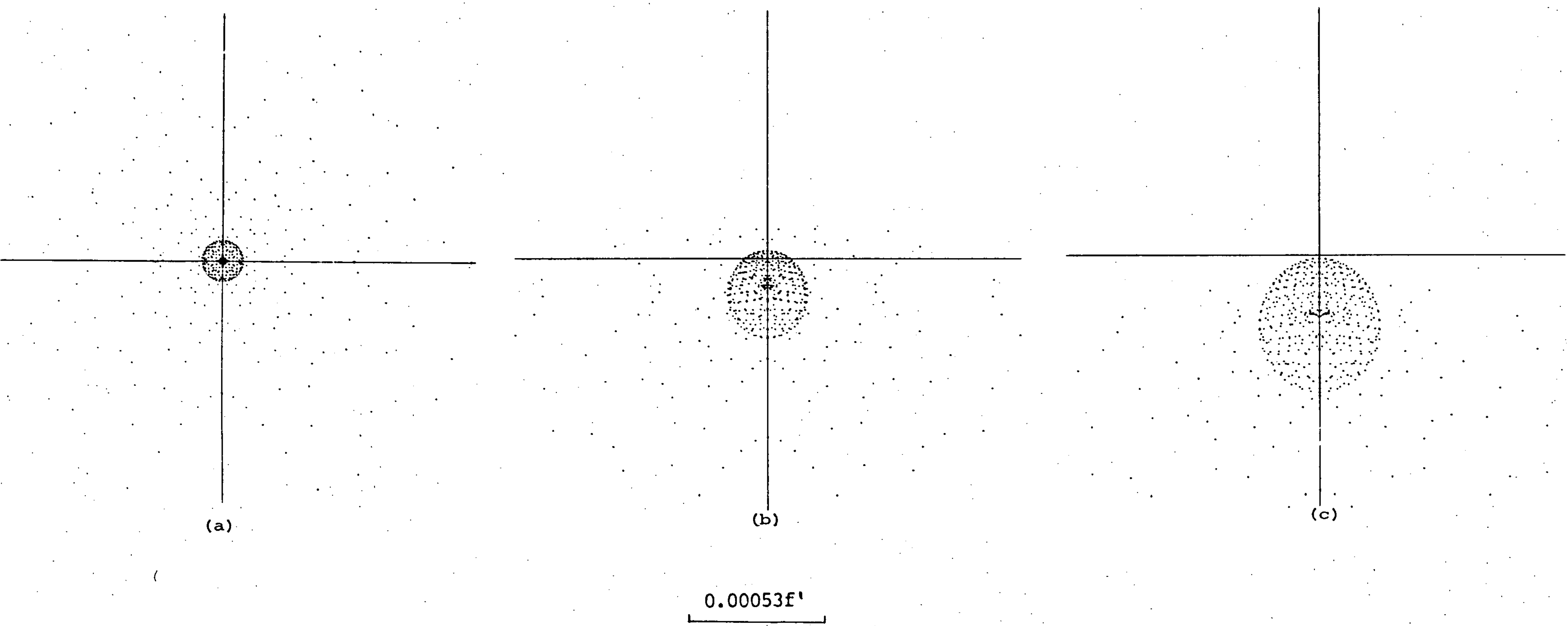


fig. 4.3 Σ_1, stage 1; a) axial, b) 5°, c) 7.5°.

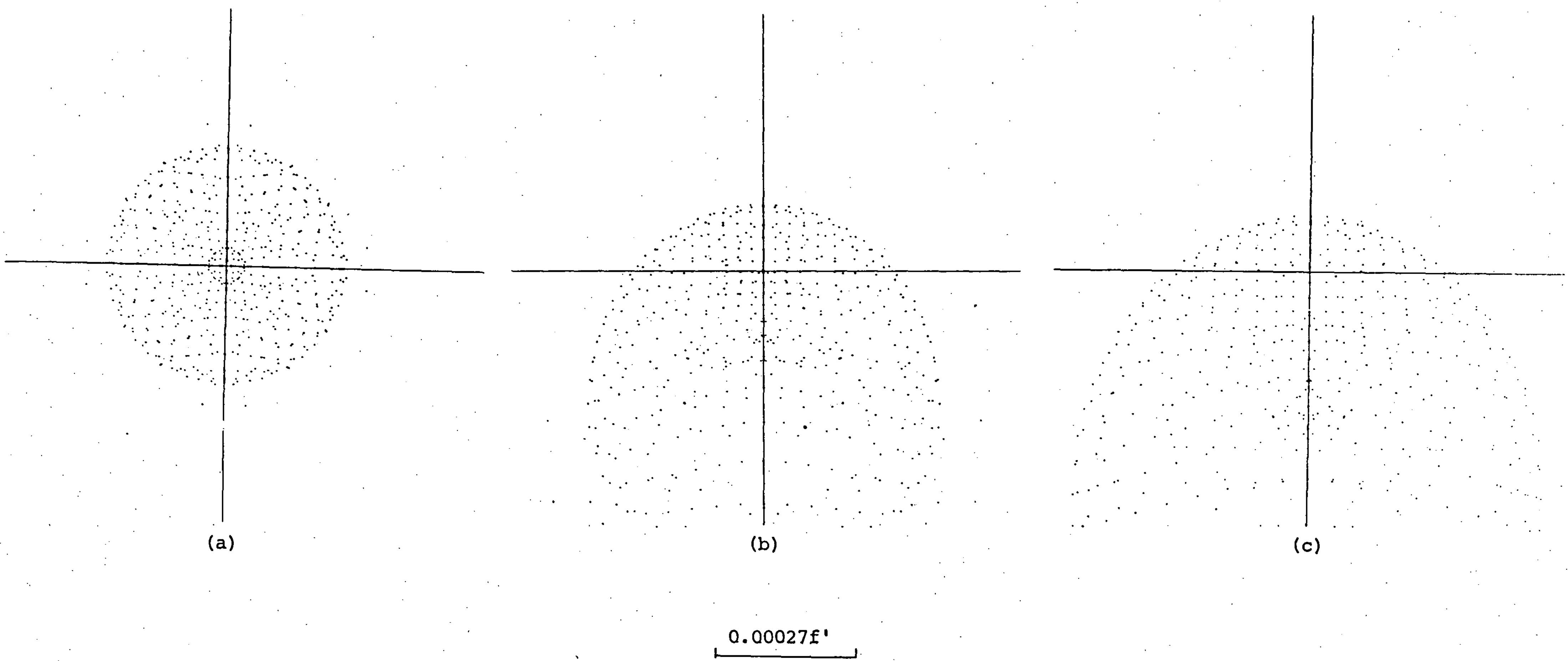


fig. 4.4 Σ_1, stage 2; a) axial, b) 5°, c) 7.5°.

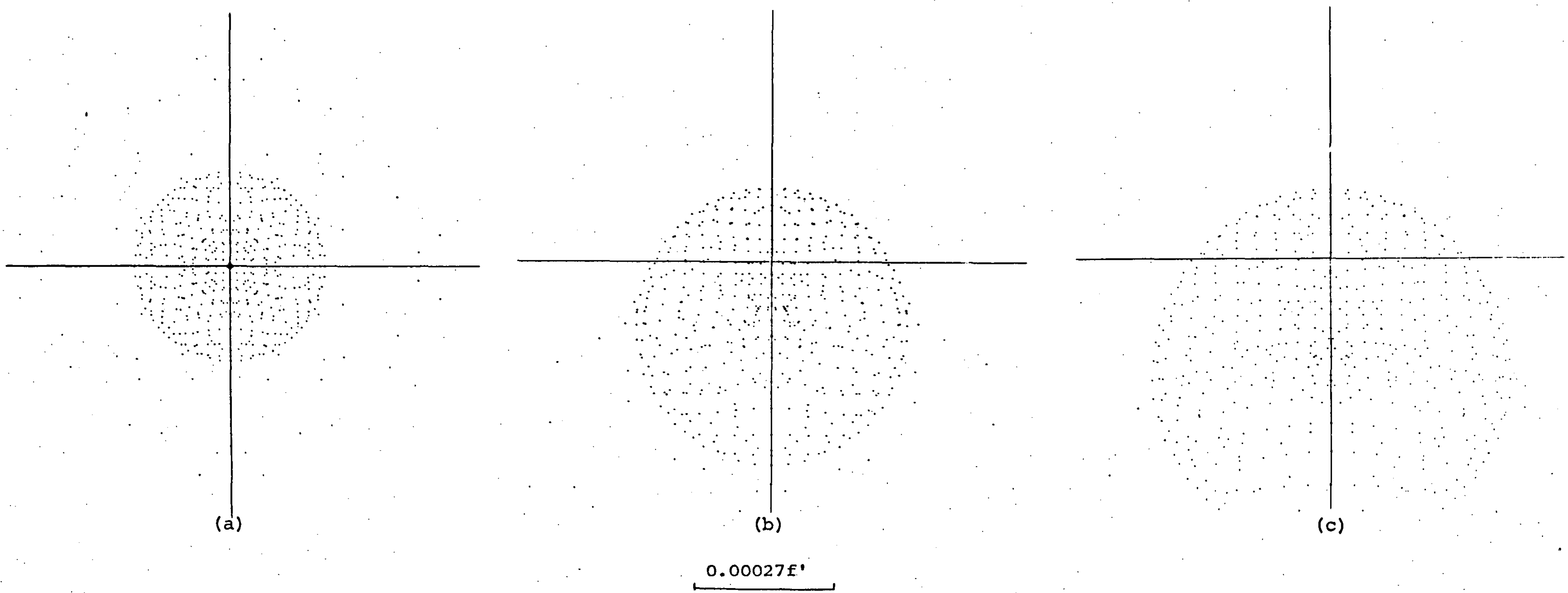


fig. 4.5 Σ_1, stage 3; a) axial, b) 5°, c) 7.5°.

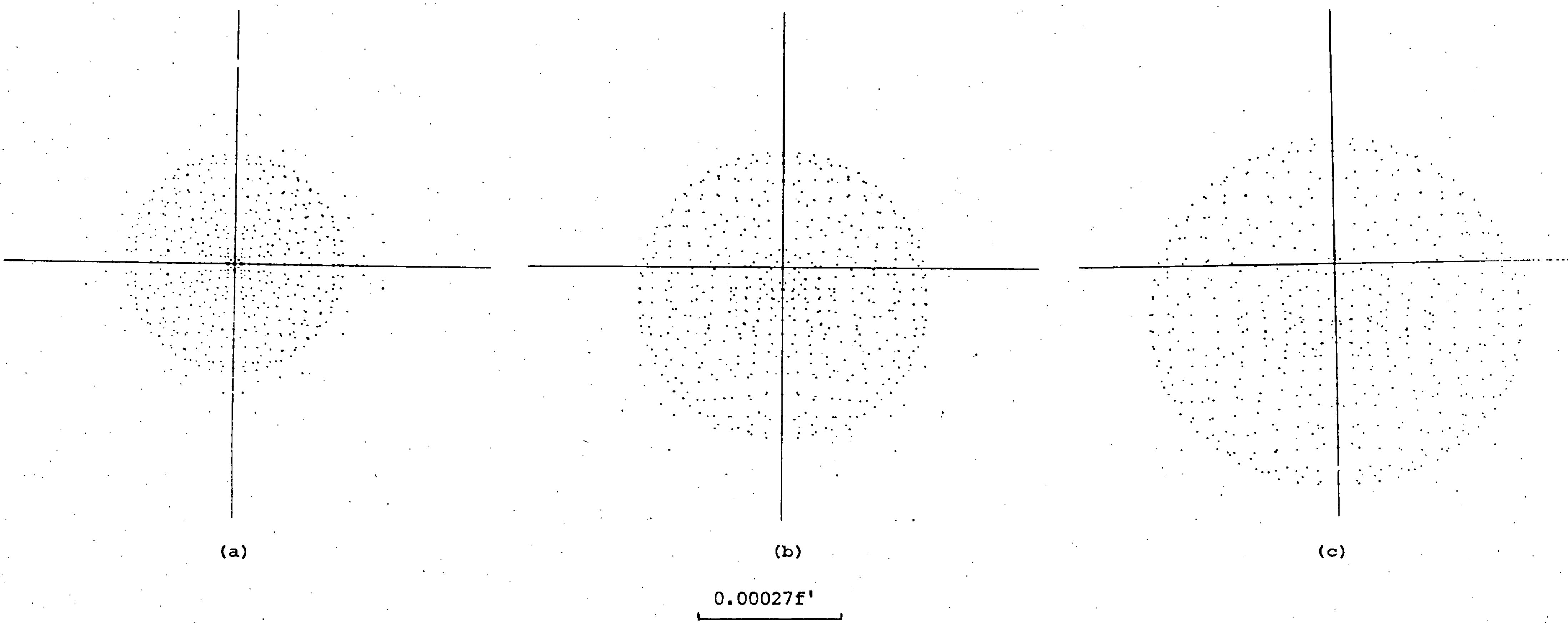


fig. 4.6 Σ_1, stage 4; a) axial, b) 5°, c) 7.5°.

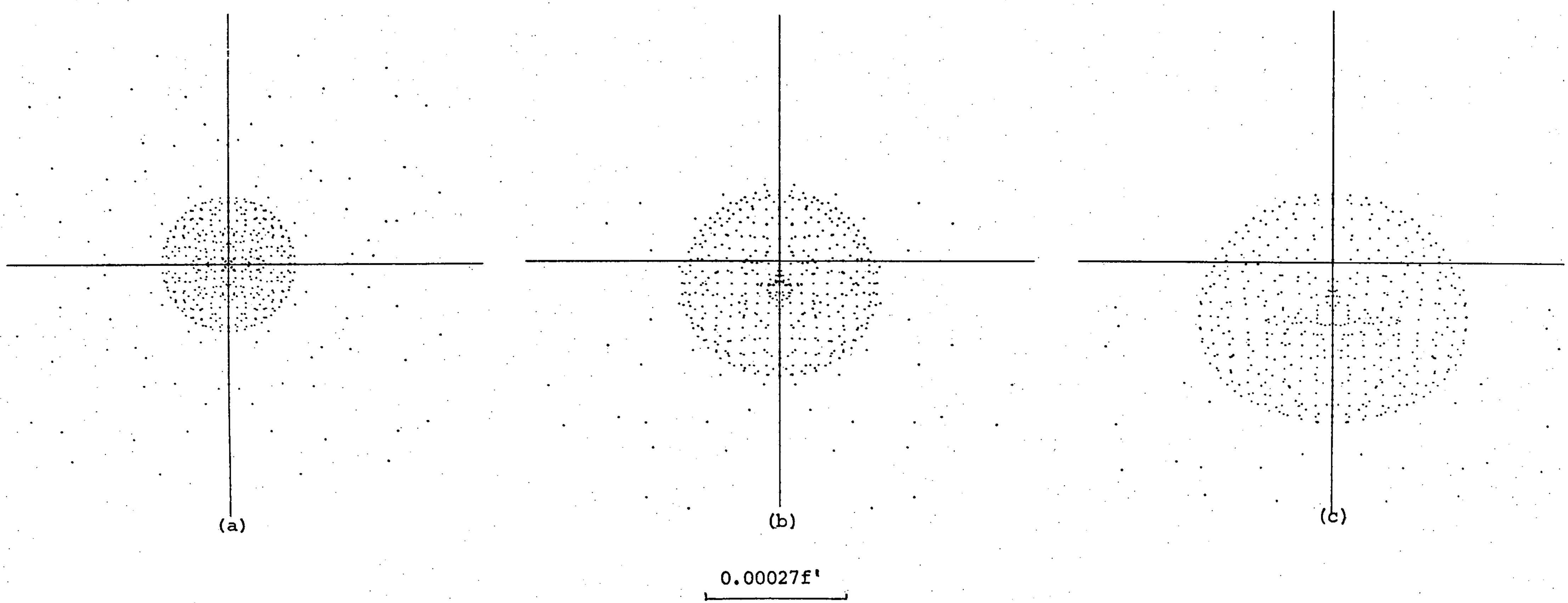


fig. 4.7 Σ_1, stage 5; a) axial, b) 5°, c) 7.5°.

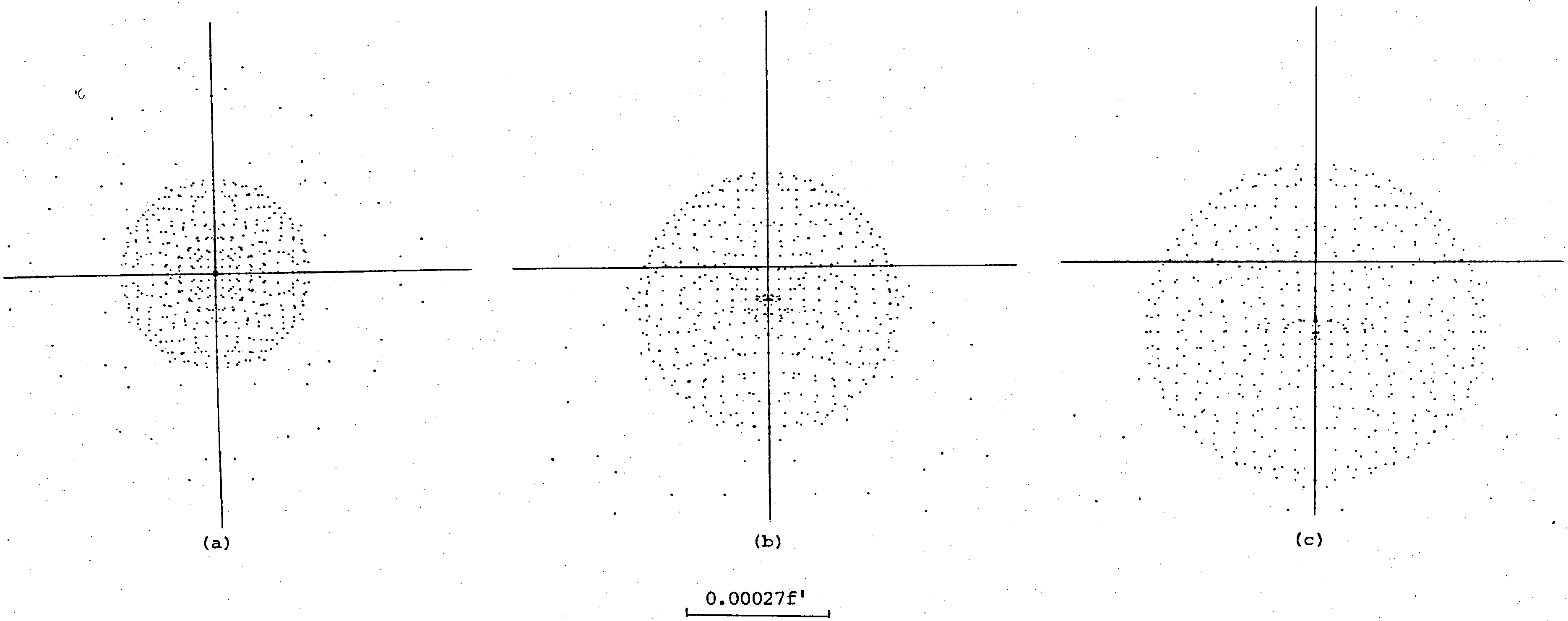


fig. 4.8 Σ_1, stage 7; a) axial, b) 5°, c) 7.5°.

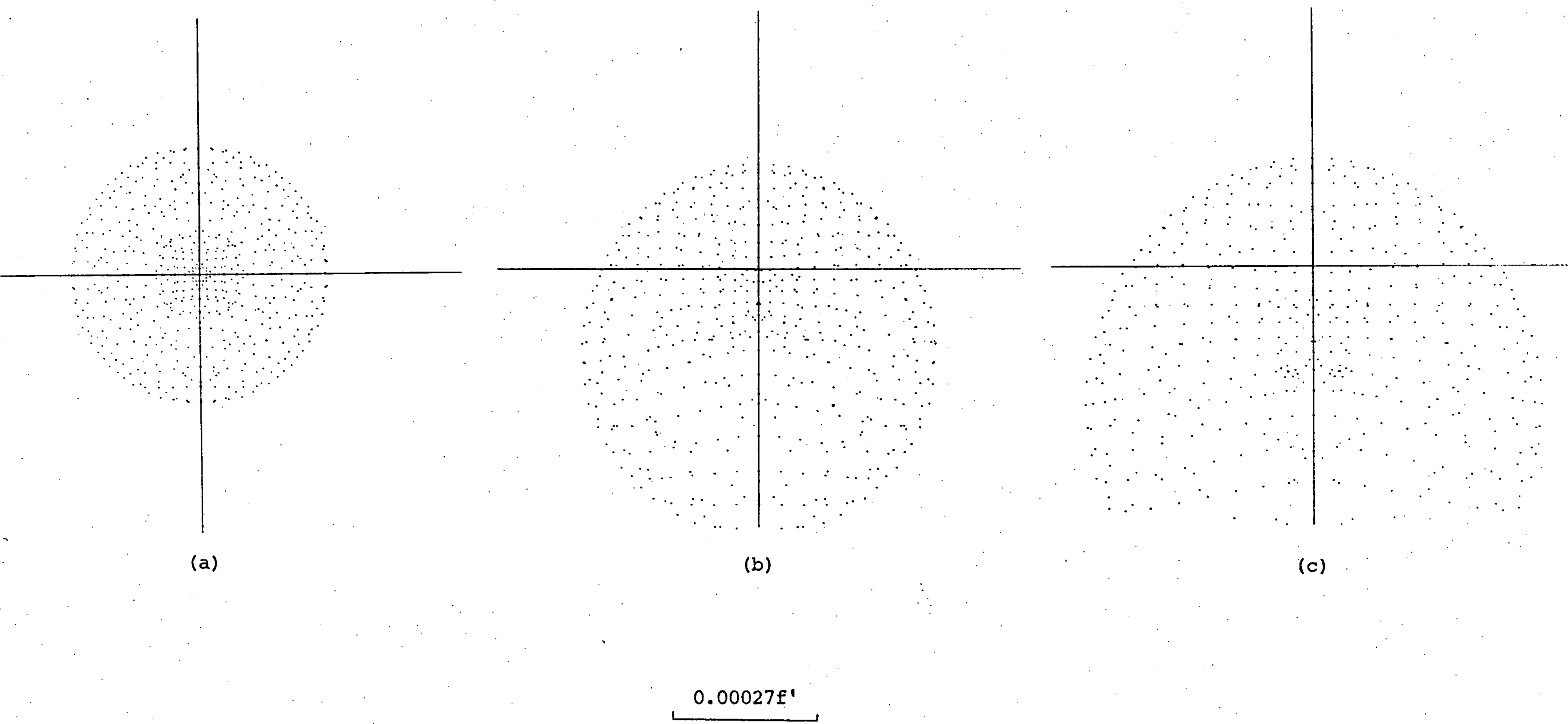


fig. 4.9 Σ_1, stage 10; a) axial, b) 5°, c) 7.5°.

These spot diagrams have been plotted for the image plan $x' = 0.0018 \times$ focal length.

§4.3 Cruickshank's Automatic Design Method for Triplet Systems

The literature on triplet design methods is quite extensive, from the work of H.D. Taylor[13] which used third order aberration theory, together with measurements of aberrations of fabricated systems, to produce satisfactory designs, through to automatic design using large computers programmed to calculate third and fifth order aberrations, perform raytracing and evaluate transfer functions, as used by R.E. Hopkins[14]. A fairly recent and extensive study of triplet design has been carried out by Aldersey[15] using an interpolative method, based largely on the spherical aberration coefficients, to choose the region within which a true optimum will lie.

The work of Cruickshank on the simple I$\dot{\text{I}}$I triplet uses paraxial conditions on the design (including colour correction) to establish a third degree equation. Choosing a particular solution of this equation allows examination of normal forms of the triplet objective. The lenses are then thickened and the third order $\bar{\text{O}}$T coefficients calculated. The third order field curvature term, fp_3+p_4 $(1 < f < 3)$, and the coma and distortion coefficients p_2 and p_5 are then adjusted to specified residuals R_3, R_2 and R_5 respectively, by changes in lens shapes. The combined power of the first two components is then used to adjust the spherical aberration (found by raytrace) to the required value R_1 at a specified aperture. This is the first step in direct control of higher order aberrations and presumably brings the system parameters into that region of parameter space in which the spherical aberration coefficients may be found to have approximately simultaneous minima (c.f. Aldersey, ref. 14, page 3). Following this longitudinal chromatic aberration is reduced to the value R_6 at a prescribed aperture, and transverse chromatic aberration to R_7 at a prescribed field angle. The fifth order coefficients are then calculated, and the residual distortion found by raytrace. The values of R_2 and R_5 are then modified, R_2 being chosen so that third order coma balances some part of the fifth order coma, and third order distortion balances higher order distortion. The flow diagram of fig. 4.10 summarizes the optimization process.

The process outlined here assumes that glass parameters and lens thicknesses have been specified. However it must be pointed out that Cruickshank and co-workers have made extensive use of the concept of fictitious glasses during much of their design studies, and currently the triplet design process propounded involves a two stage process - the first assumes a continuous range of glass parameters and controls the Petzval coefficient, σ_4, as well as the terms specified above. The second stage is that just outlined, in which the choice of glass parameters is explicit, and will be such that σ_4 remains close to the originally specified value. By using fictitious glasses having a continuous range of refractive index and V-number within certain bounds, it is possible to simulate more complex systems derived from the triplet.

The optimization process was examined during the design of a 24 inch focal length objective having an aperture of F/6 and covering a 10° semi-field. This corresponds approximately to the specifications used by R.E. Hopkins[14] for some near-optimum designs. The glasses chosen were from the Schott catalogue, LaK10/SF10/LaK10. Since only the monochromatic design process was being examined those parts of the program used in chromatic correction were omitted. The progress of the design was followed by examining the system after each change of the thin lens parameter χ, or the residuals R_2 and R_5. By choosing only these attention is confined to those changes involved in the correction of higher orders of aberration. In figs. 4.11 to 4.19 are shown spot diagrams calculated at 0°, 7° and 10° for various stages of the design process. These are given for the gaussian image plane, which, in this case, is close to being the optimum image plane.

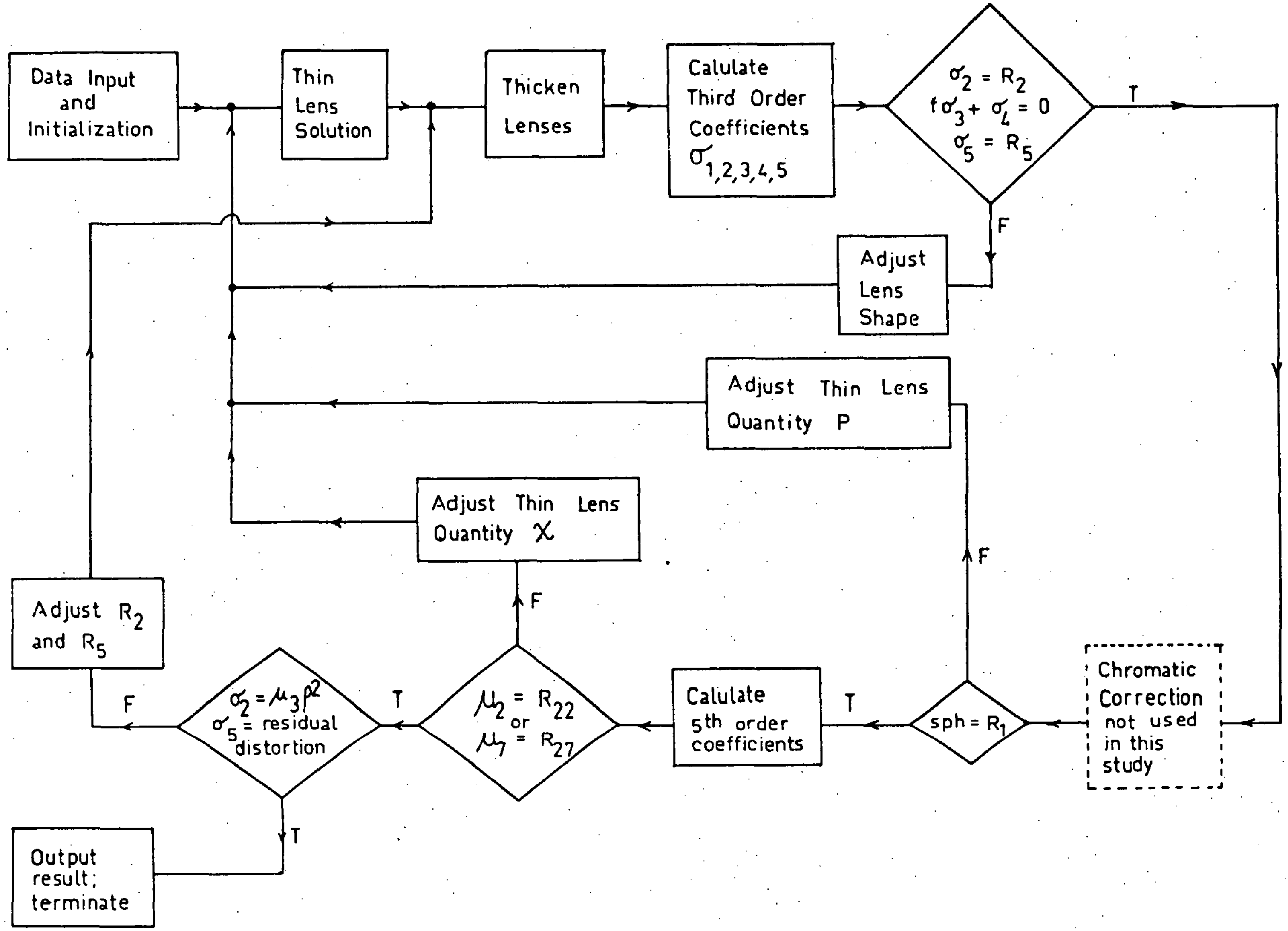

Data Input and Initialization
Thin Lens Solution
Thicken Lenses
Calulate Third Order Coefficients $\sigma_{1,2,3,4,5}$
$\sigma_2 = R_2$ $f\sigma_3 + \sigma_4 = 0$ $\sigma_5 = R_5$
T
F
Adjust Lens Shape
Adjust Thin Lens Quantity P
Adjust Thin Lens Quantity χ
F
F
Adjust R_2 and R_5
F
$\sigma_2 = \mu_3 \rho^2$ σ_5 = residual distortion
T
$\mu_2 = R_{22}$ or $\mu_7 = R_{27}$
Calulate 5th order coefficients
T
sph = R_1
Chromatic Correction not used in this study
T
Output result; terminate

fig. 4.10

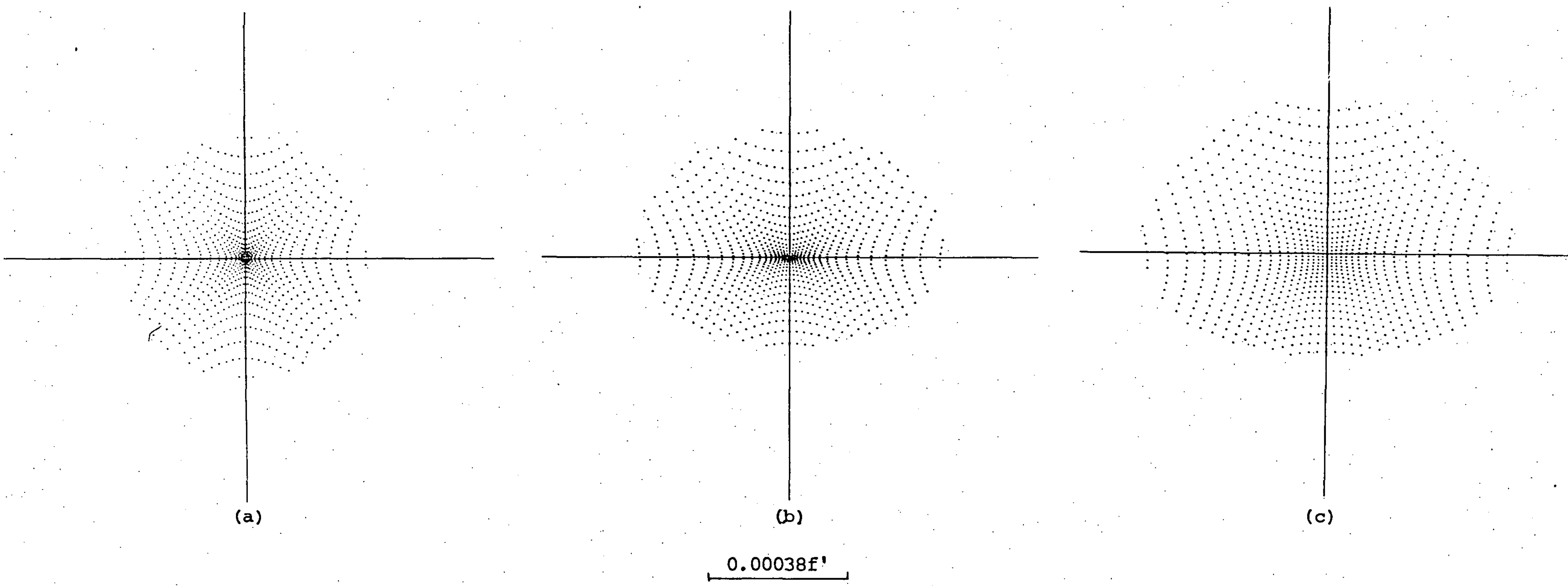


fig. 4.11 CCl, stage 1; a) axial, b) 7°, c) 10.5°.

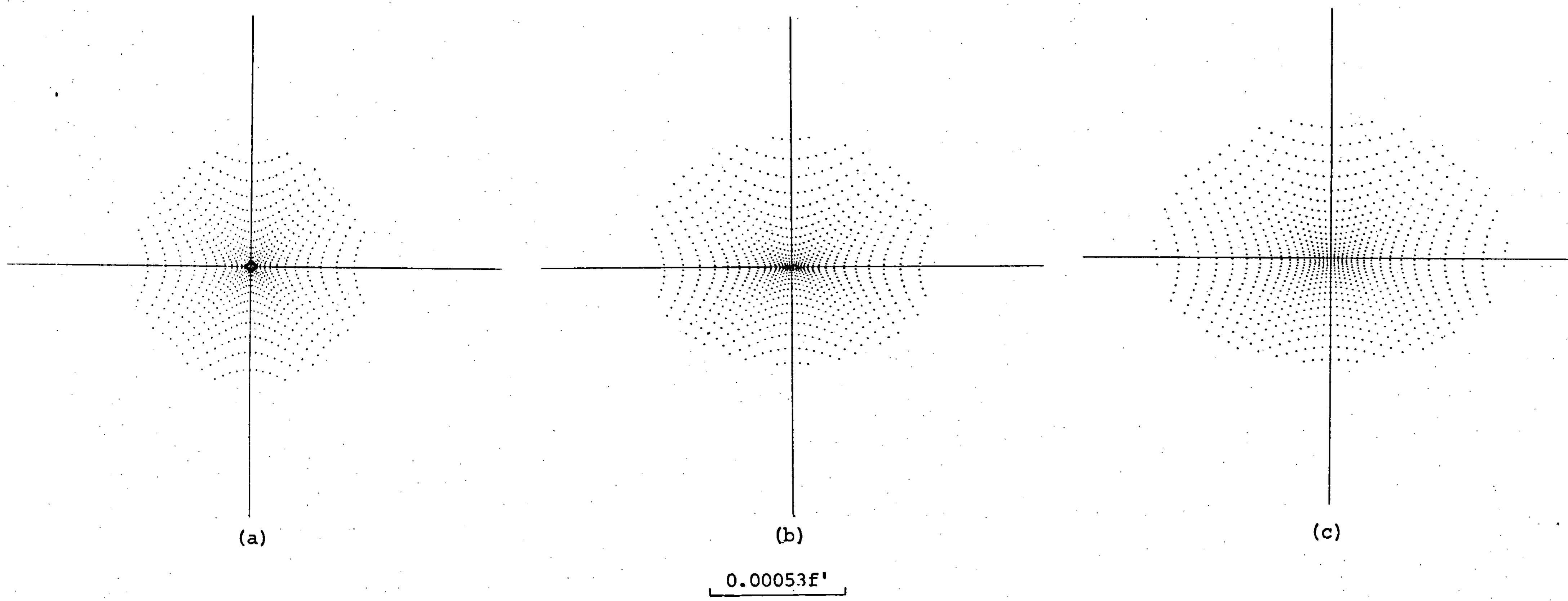


fig. 4.12 ÇCl, stage 2; a) axial, b) 7°, c) 10.5°.

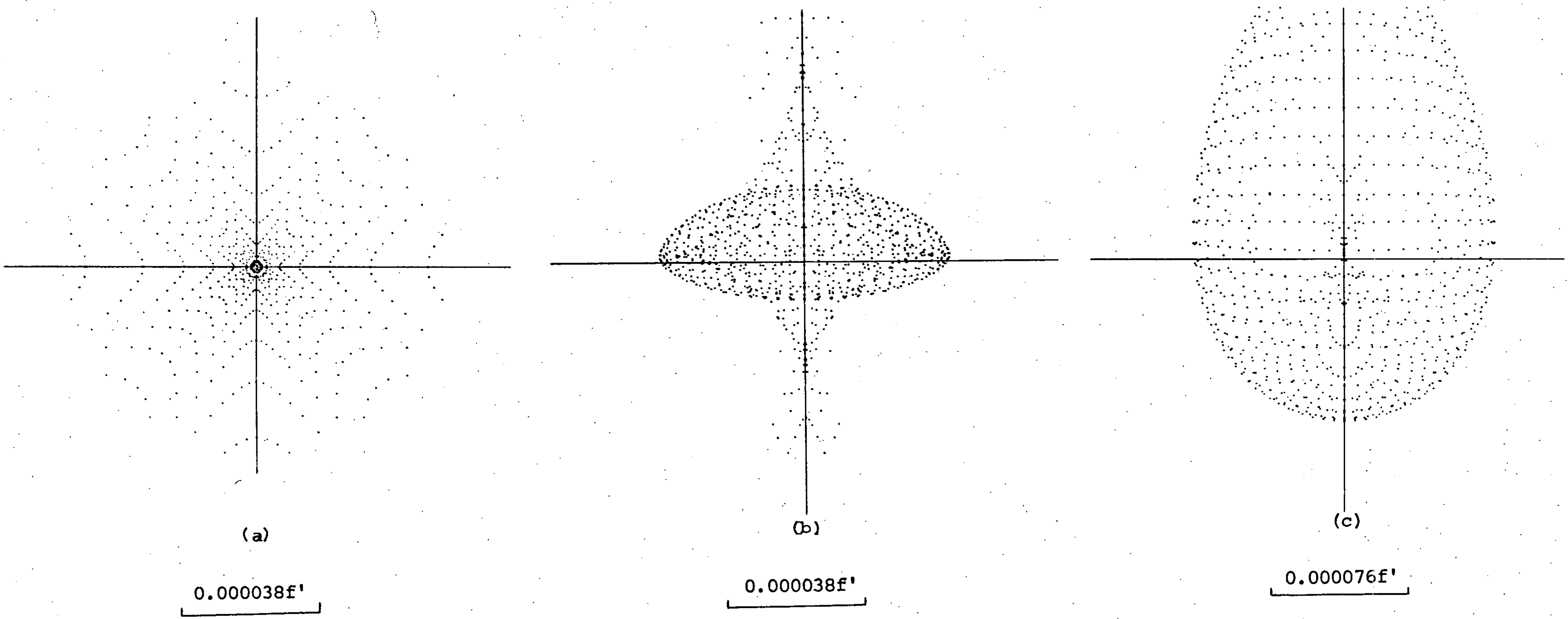


fig. 4.13 CCl, stage 4; a) axial, b) 7°, c) 10.5°.

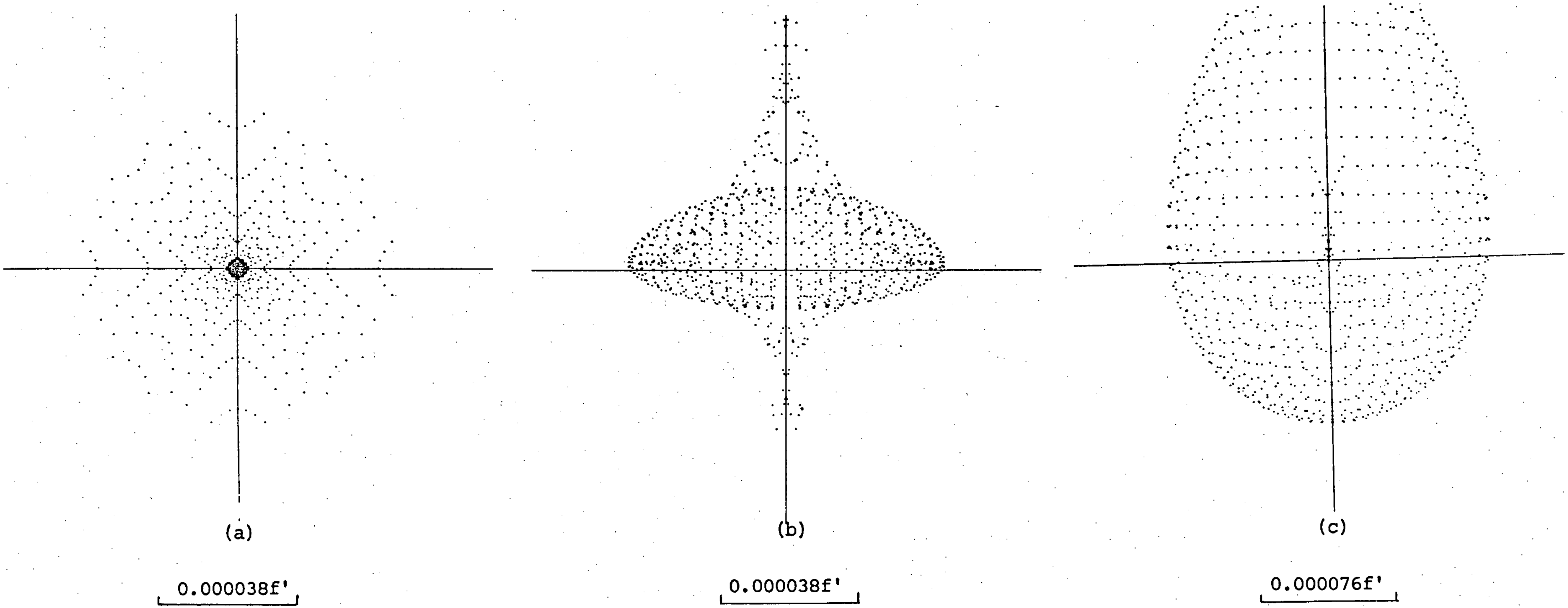


fig. 4.14 CCl, stage 5; a) axial, b) 7°, c) 10.5°.

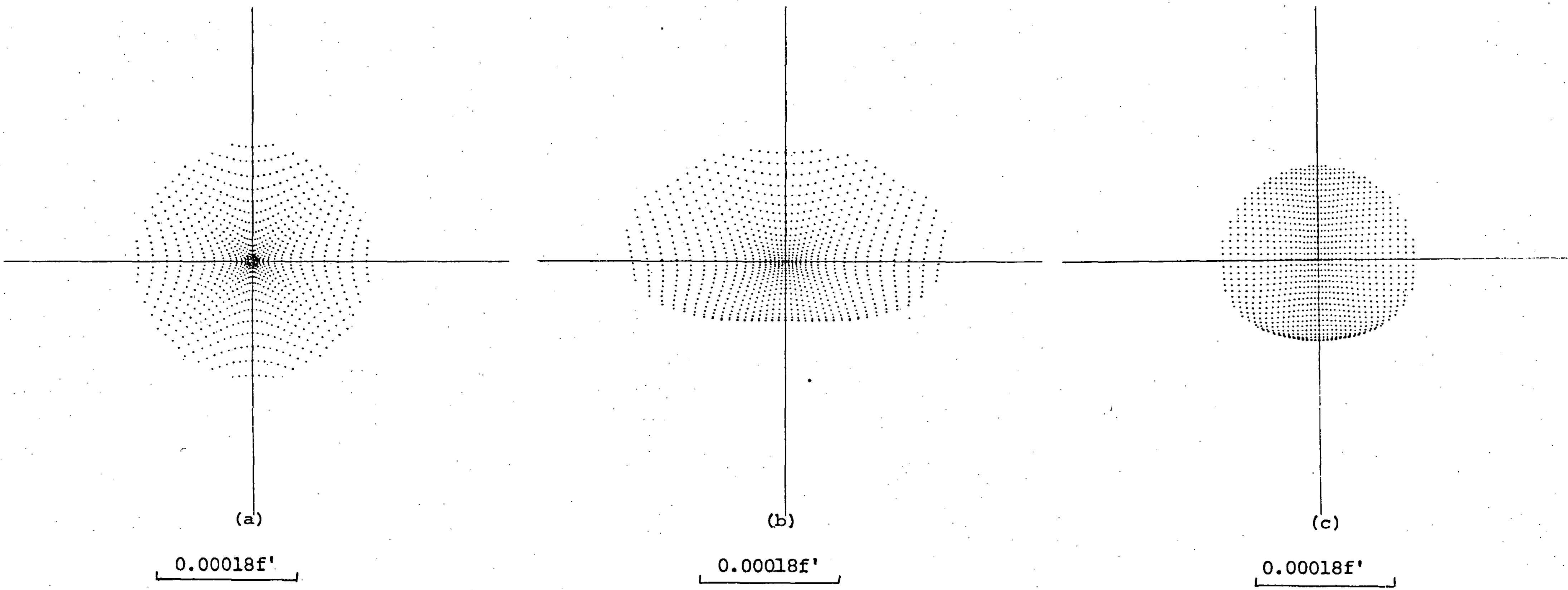


fig. 4.15 CCl, stage 6; a) axial, b) 7°, c) 10.5°.

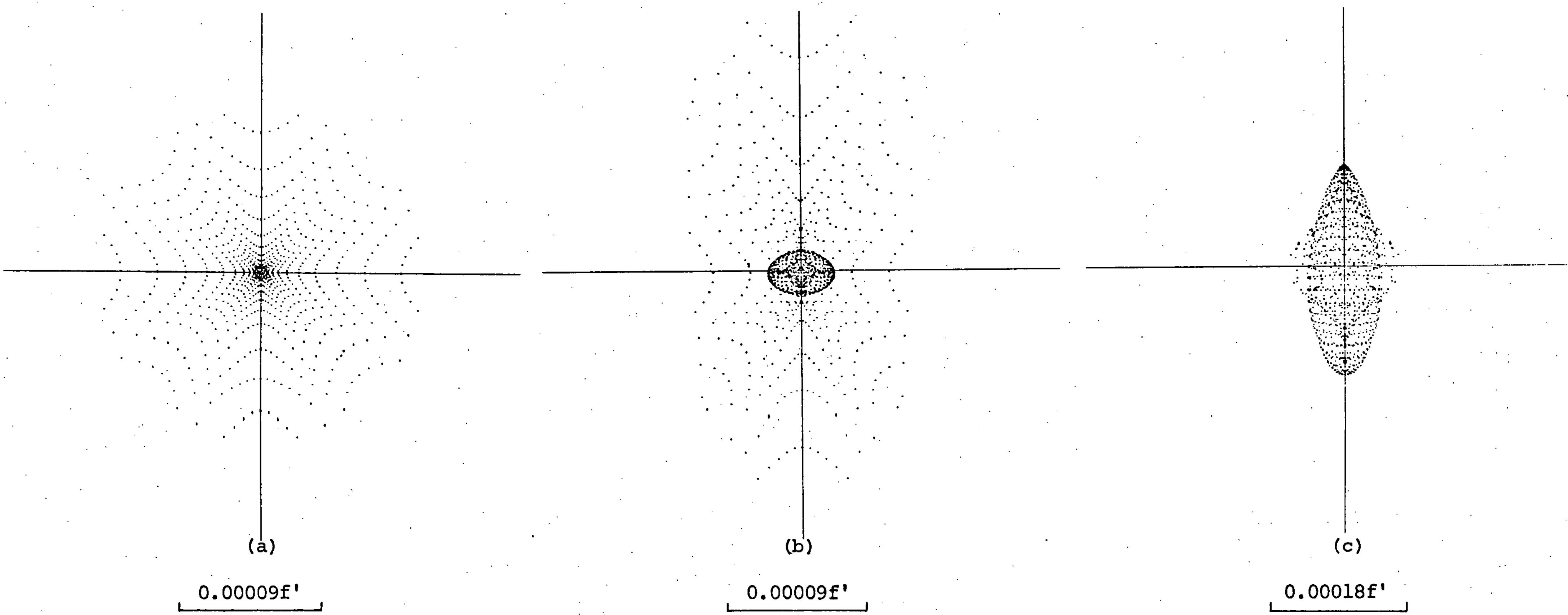


fig. 4.16 CCl, stage 7; a) axial, b) 7°, c) 10.5°.

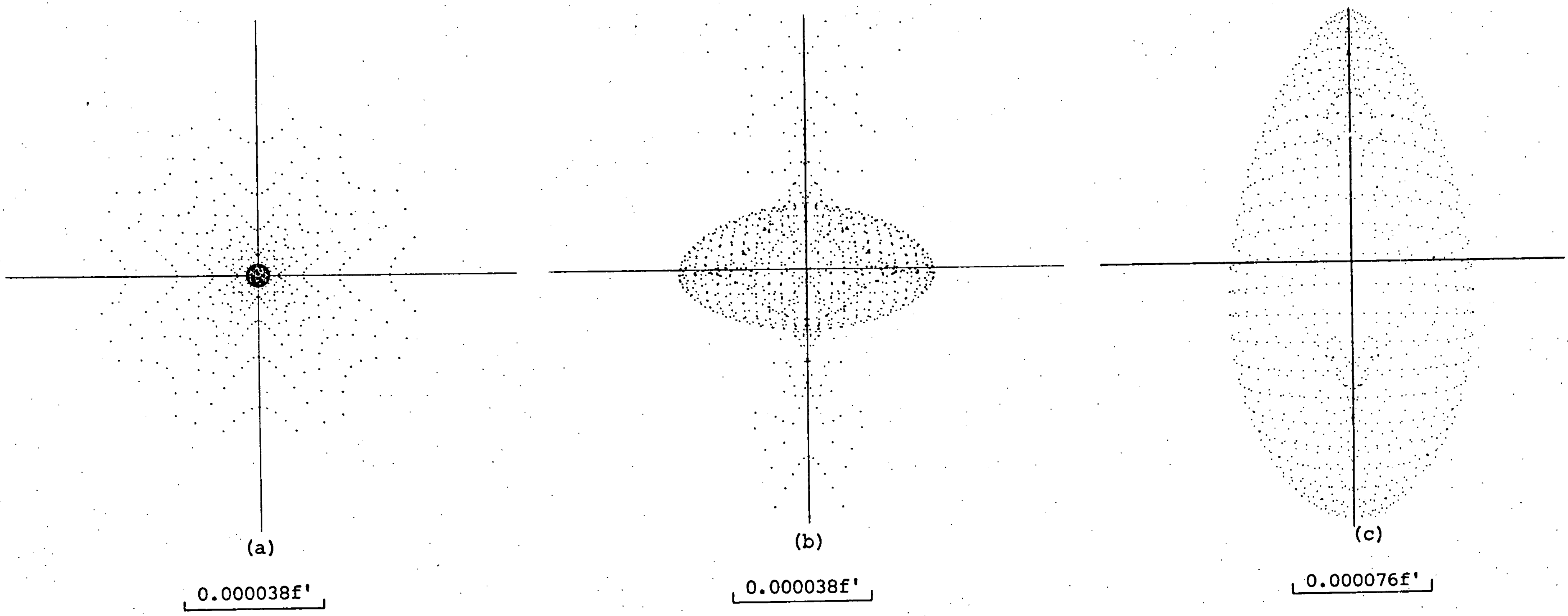


fig. 4.17 CCl, stage 8; a) axial, b) 7°, c) 10.5°.

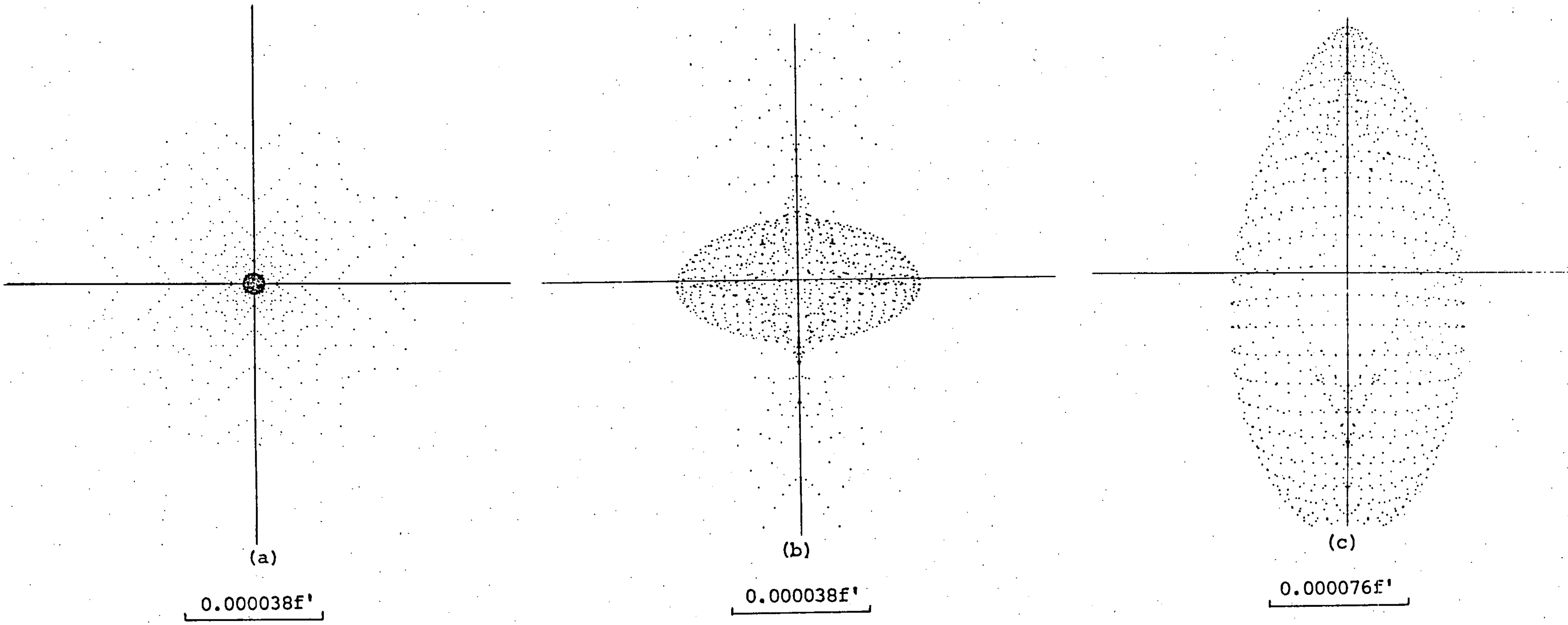


fig. 4.18 CCl, stage 10; a) axial, b) 7°, c) 10.5°.

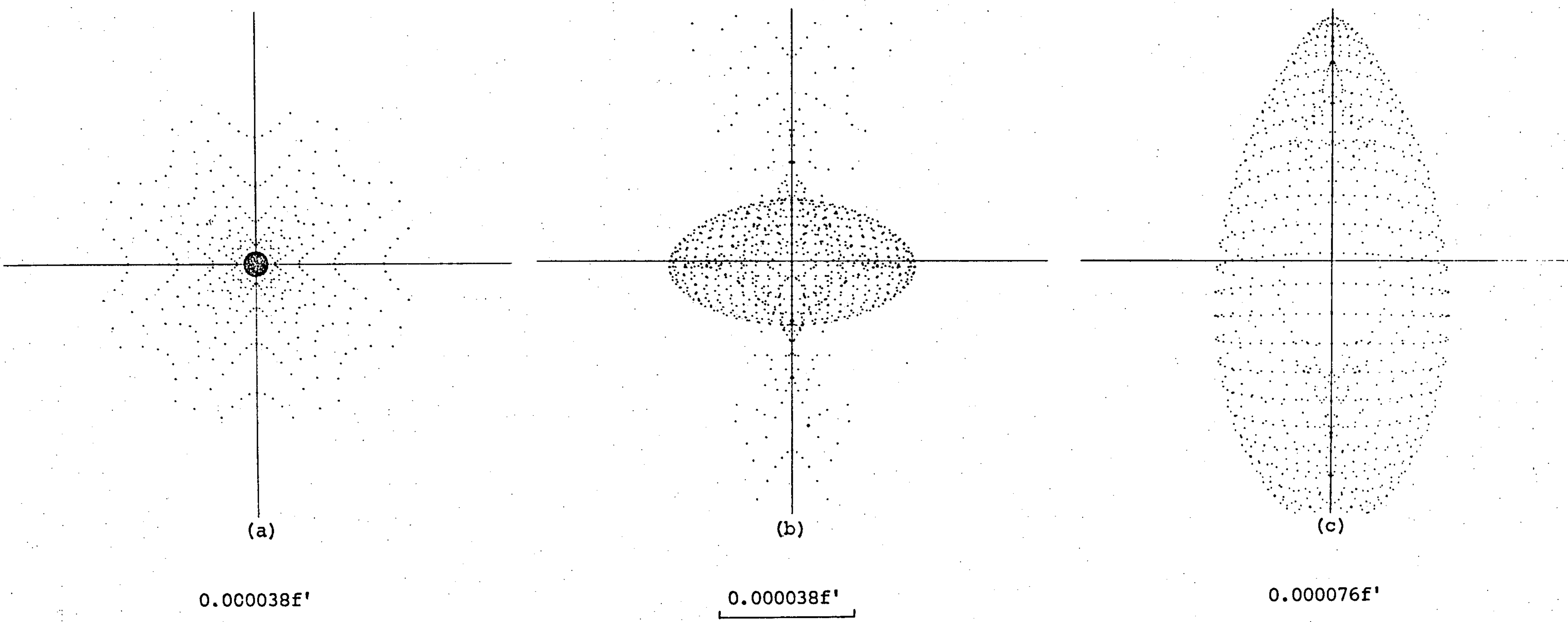


fig. 4.19 CCl, stage 11; a) axial, b) 7°, c) 10.5°.

This system has been designated CC for convenience of reference. It will be noted that the spot diagrams only continue to stage ten of the process, though in fact another thirteen such stages were completed before termination. However it was found that the changes resulting from these further iterations were quite insignificant, as indicated by the values of the assessment functions given in the following section.

§4.4 Studies of the Optimization Process using Different Criteria

In this section both the optimization processes detailed above are examined independently using the assessment functions introduced earlier. As in the previous chapter, we use the functions P_1 (radius of gyration of the spot diagram), Var (the variance of the wavefront retardation), $S_1(r)$ (the o.t.f. - based function introduced in chapter two), and $S_{1M}(r)$, (the analogue of $S_1(r)$ based on the m.t.f. and calculated using the variance of the aberration difference function), to measure image spread. The functions $P_2(=\overline{\varepsilon_y^2}-\overline{\varepsilon_z^2})$, $S_2(r)$, $(\mathrm{Re}\{T(r,\pi/2)-T(r,o\})$, and $S_{2M}(r)(=V(r,\pi/2)-V(r,o))$ are used to measure astigmatism.

a) Optimization using Spot Diagram Functions

The graphs of the values of various assessment functions at successive stages of the optimization are shown in figs. 4.20 to 4.37. Perhaps most important to note is that P_1 is decreased considerably during the process, particularly in the first four steps. Thereafter P_1 is constrained while P_2 is reduced to near zero (steps 5,6,7 and 8). Attempts to reduce P_3 led to failure of the process. However during the initial eight steps P_3 was being systematically reduced as required. Hence the optimization process is successful under some conditions, but has been found to fail due to there being no solution to the linear programming problem under the imposed constraints. It should be re-emphasized that the writer does not consider

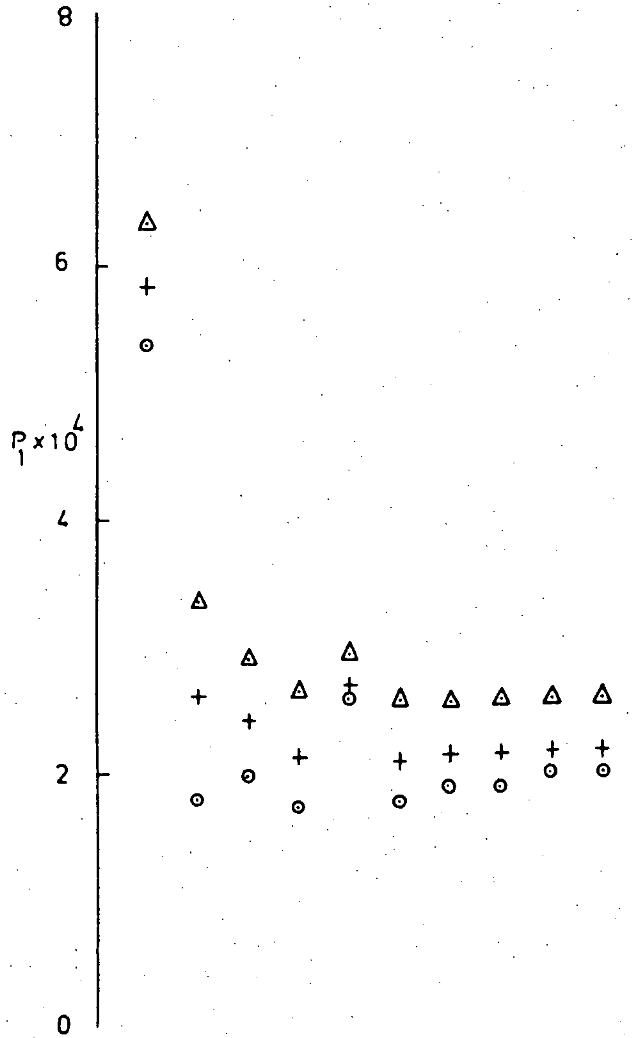


fig.4.20 Radius of gyration

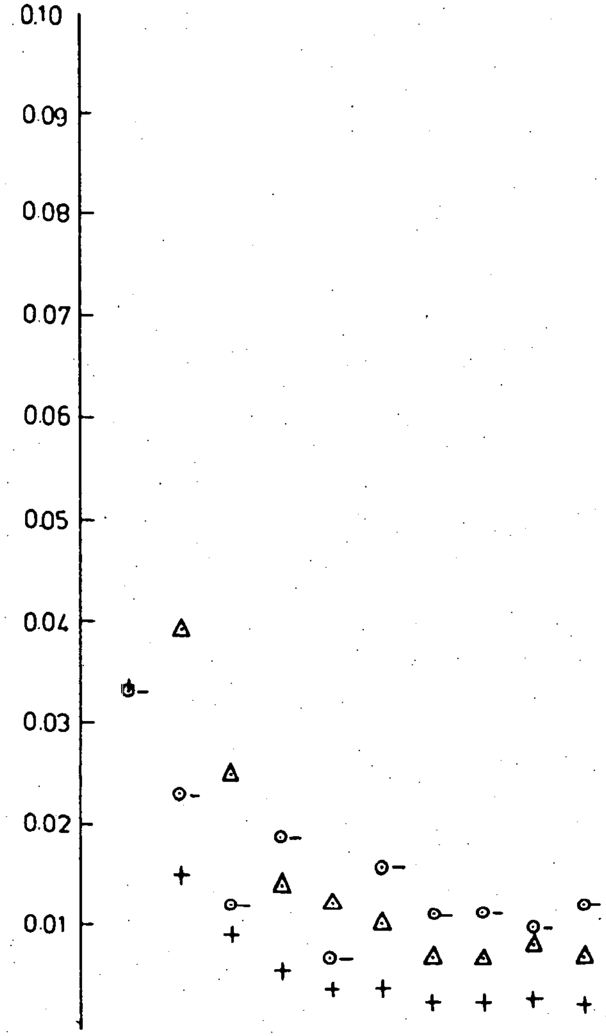


fig.4.21 Wavefront Variance

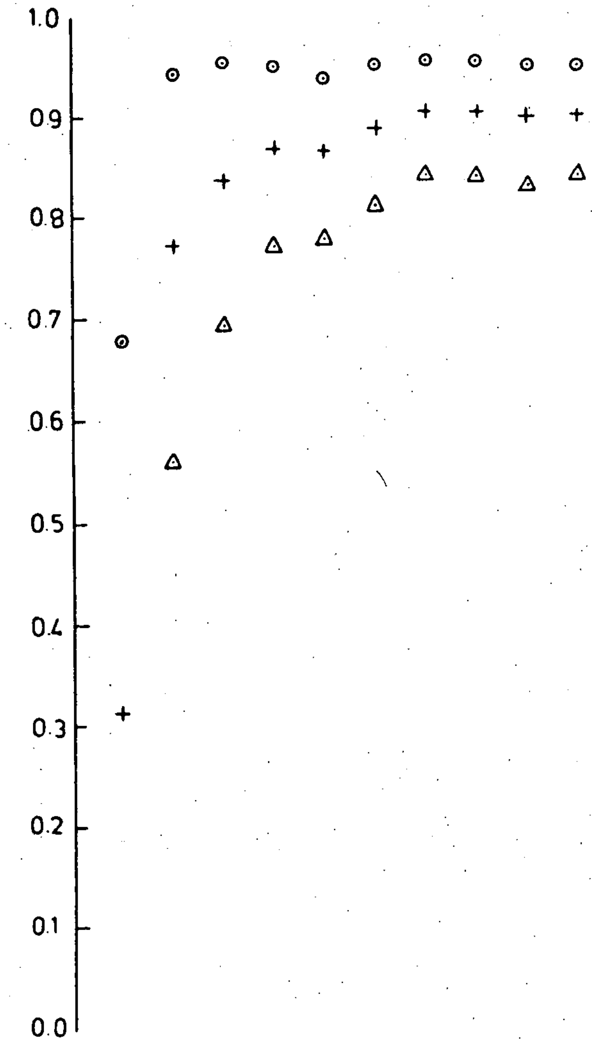


fig.4.22 $S_{1M}(0.057)$

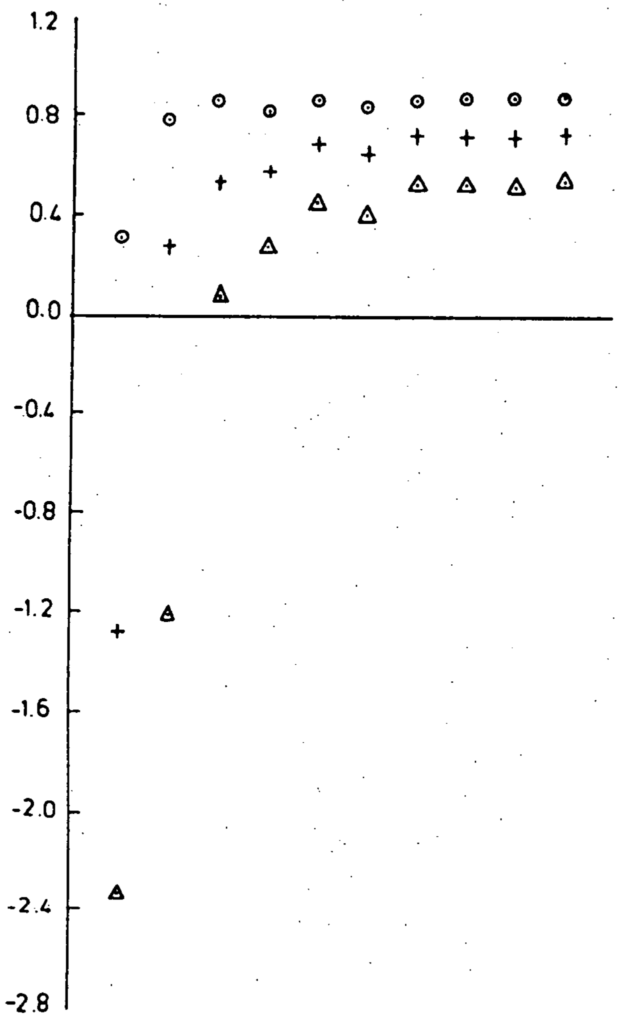


fig. 4.23 $S_{1M}(0.114)$

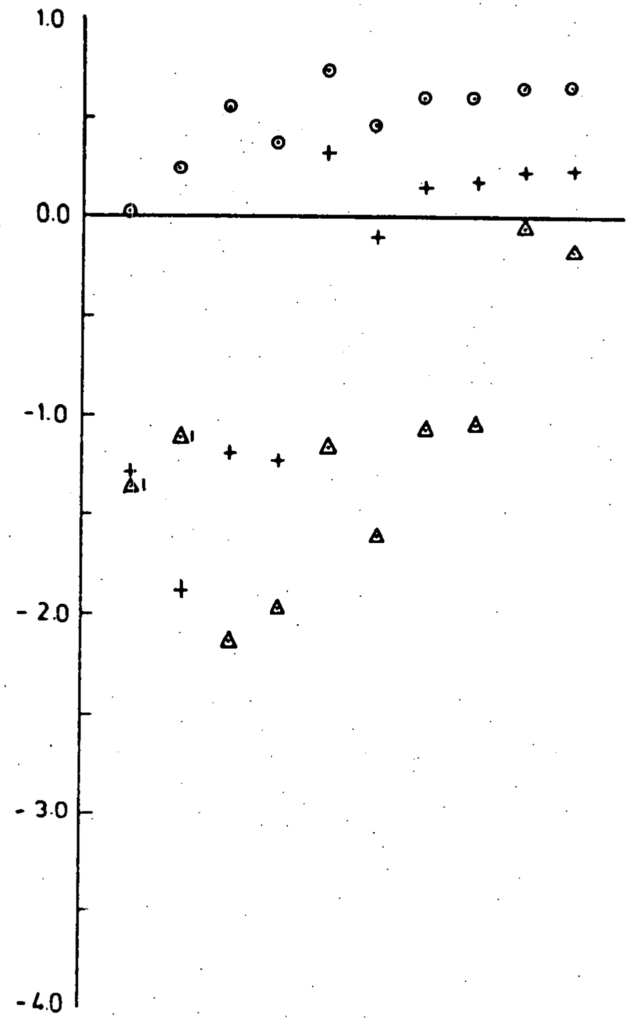


fig. 4.24 $S_{1M}(0.228)$

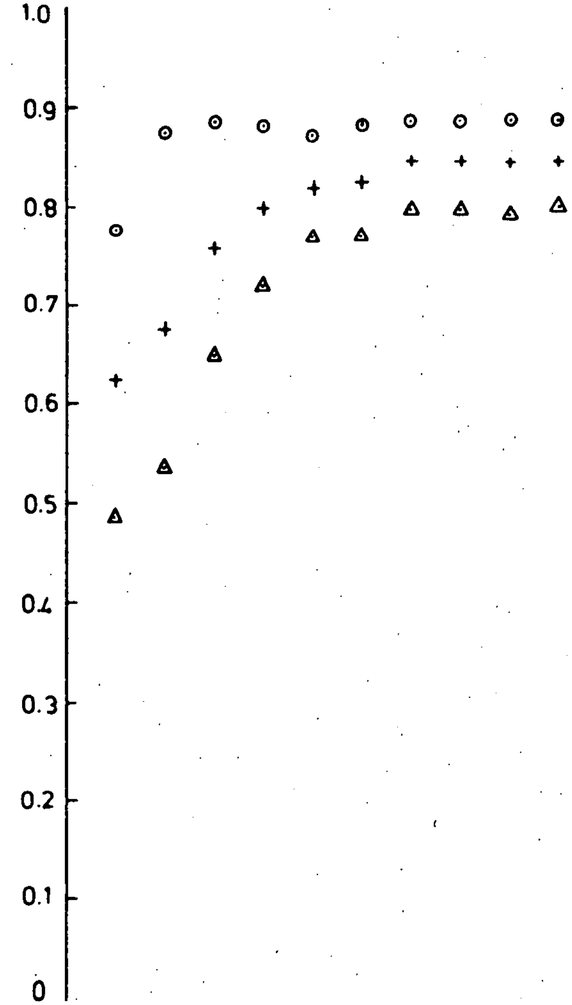


fig. 4.25 $S_{1}(0.057)$

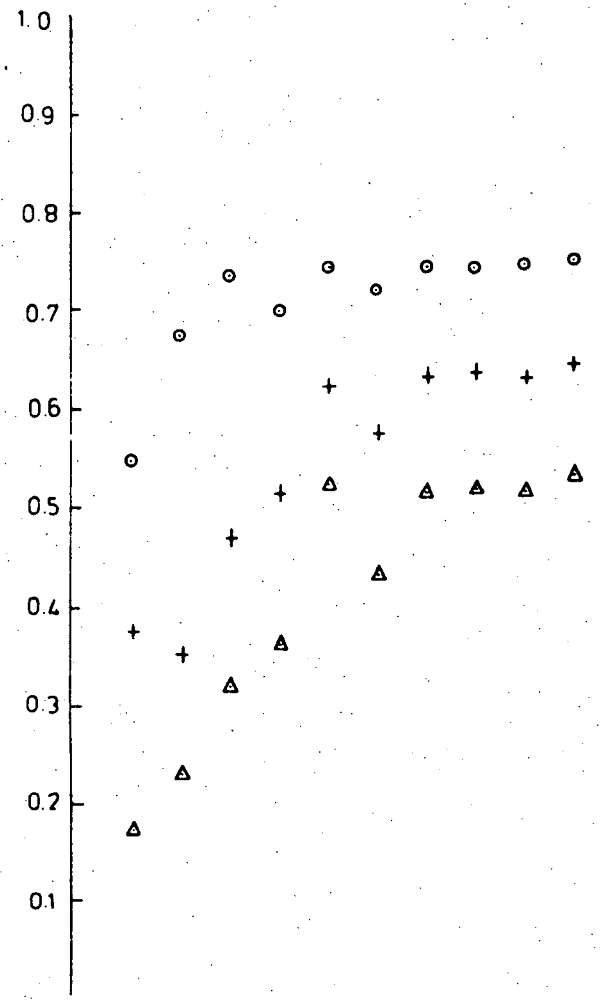


fig. 4.26 $S_1(0.114)$

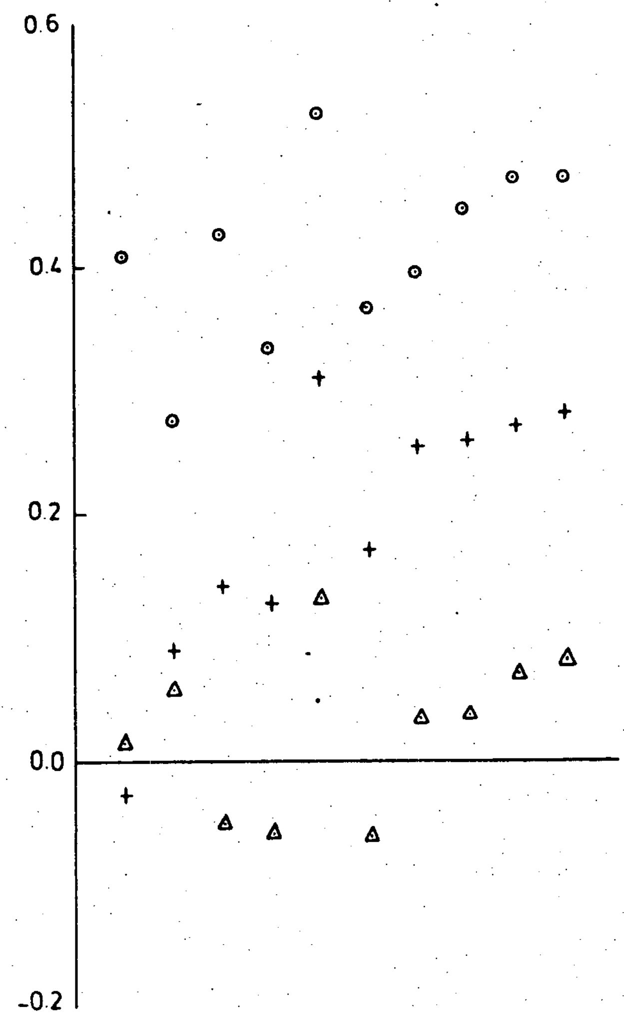


fig. 4.27 $S_1(0.228)$

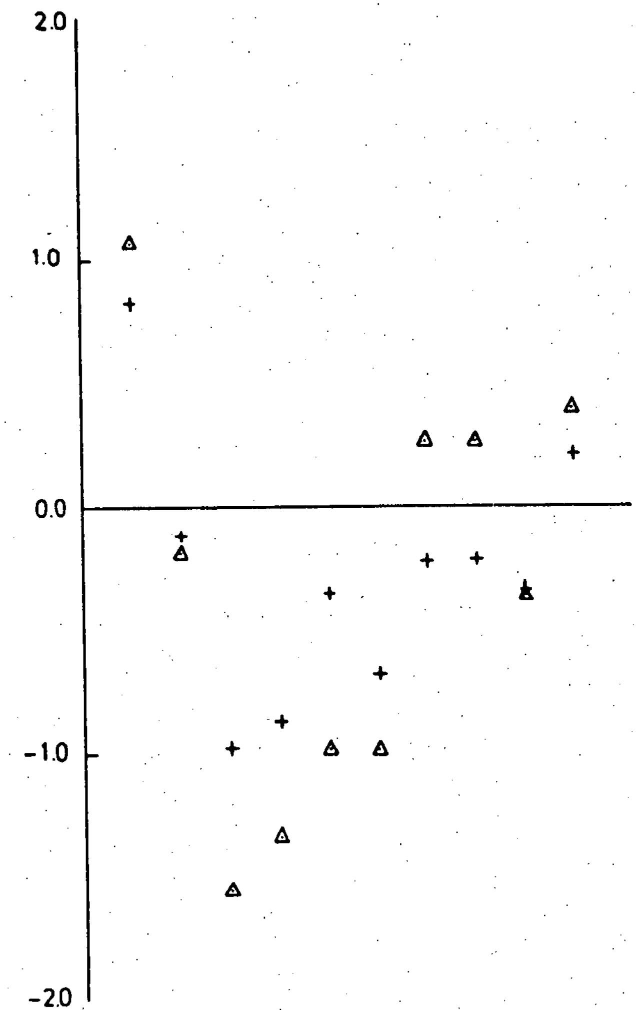


fig. 4.28 $P_2 \times 10^4$

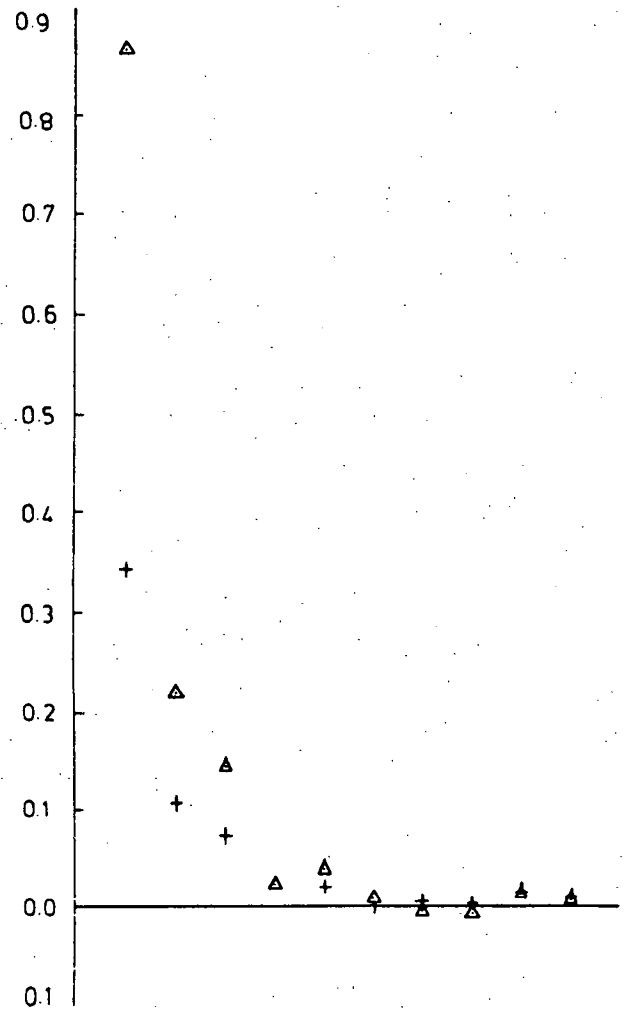


fig. 4.29 $S_{2M}(0.057)$

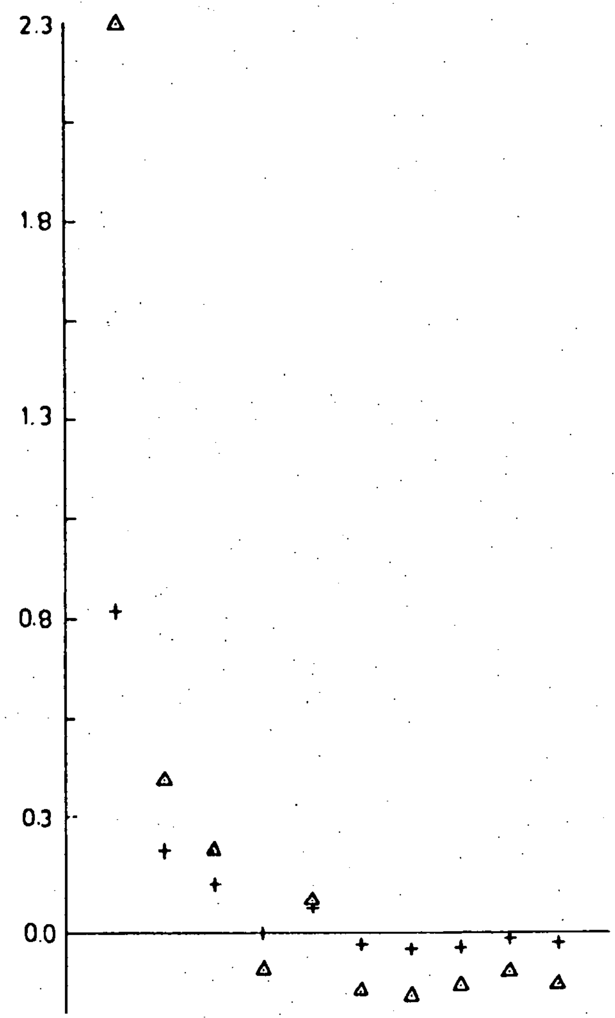


fig. 4.30 $S_{2M}(0.114)$

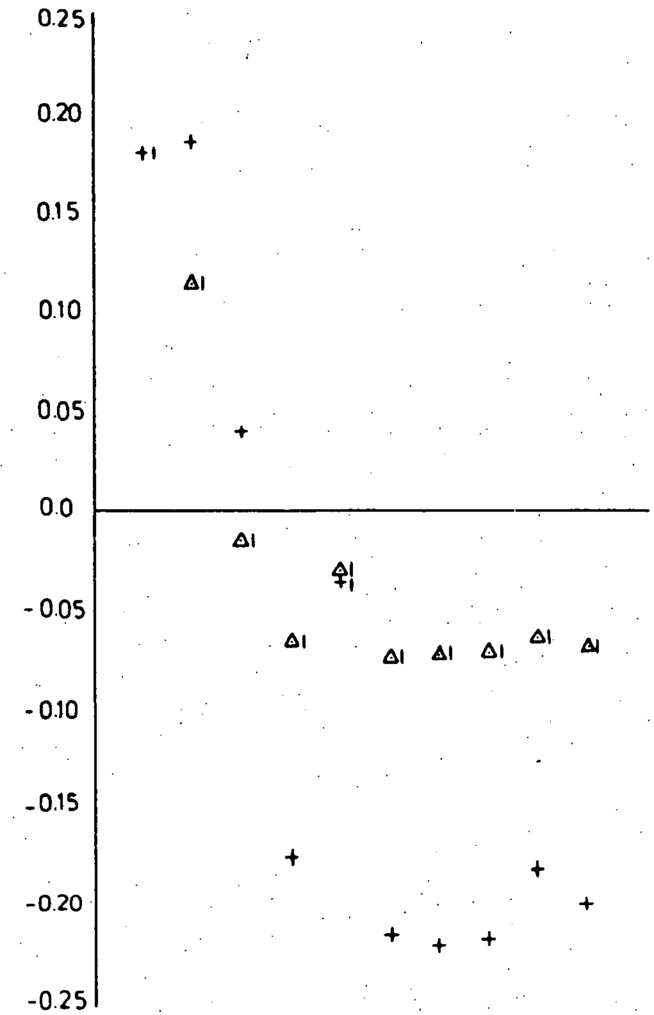


fig. 4.31 $S_{2M}(0.228)$

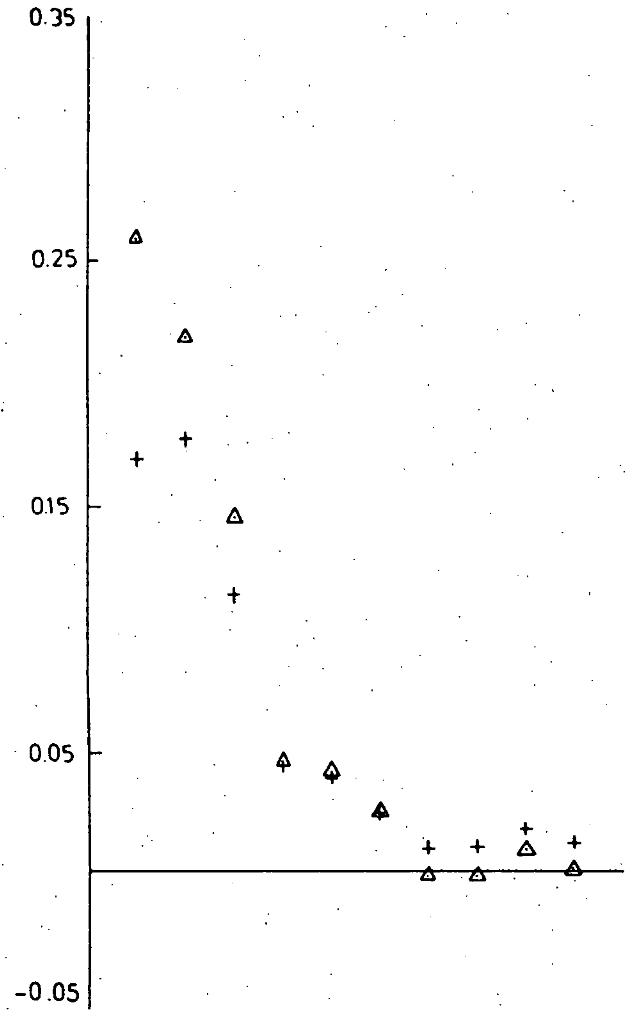


fig. 4.32 $S_2(0.057)$

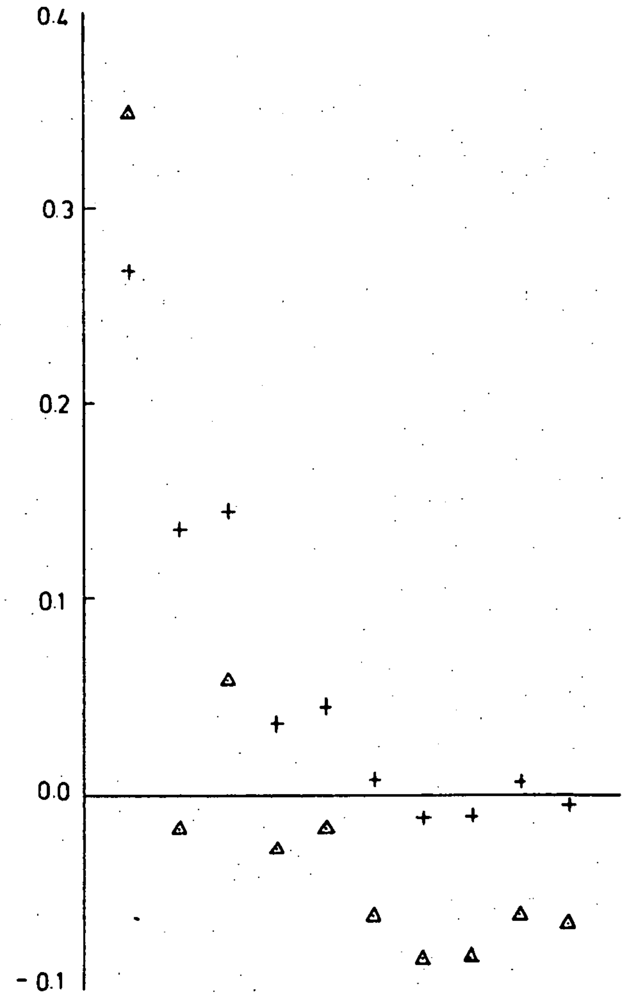


fig. 4.33 $S_2(0.114)$

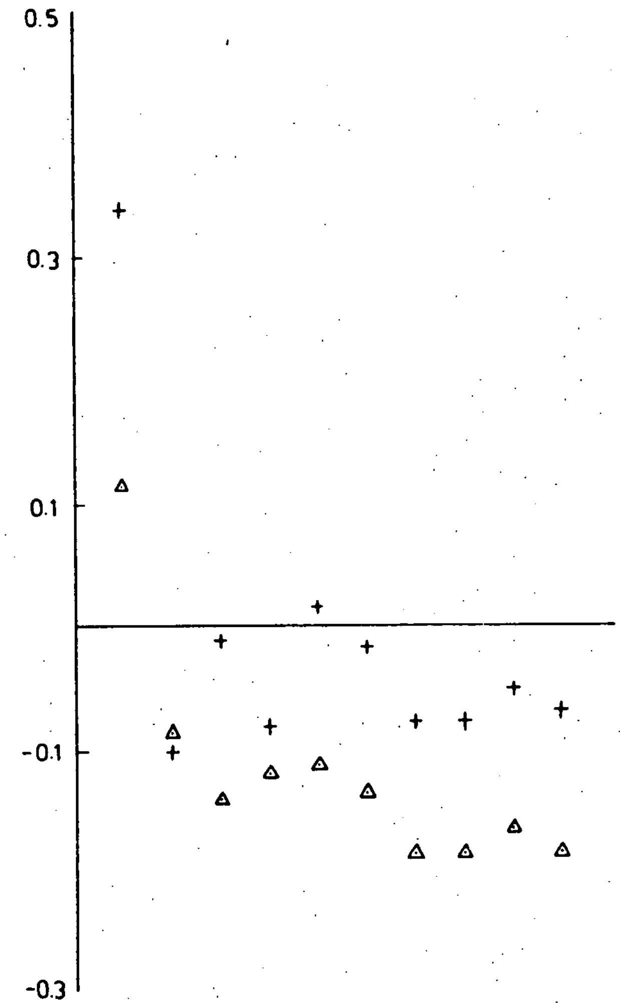


fig. 4.34 $S_2(0.228)$

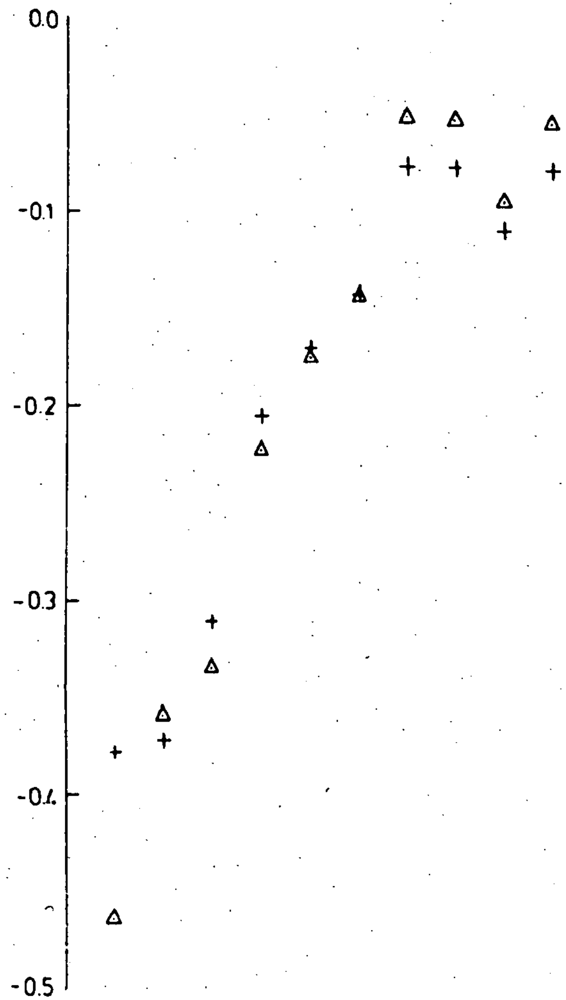


fig. 4.35 $S_3(0.057)$

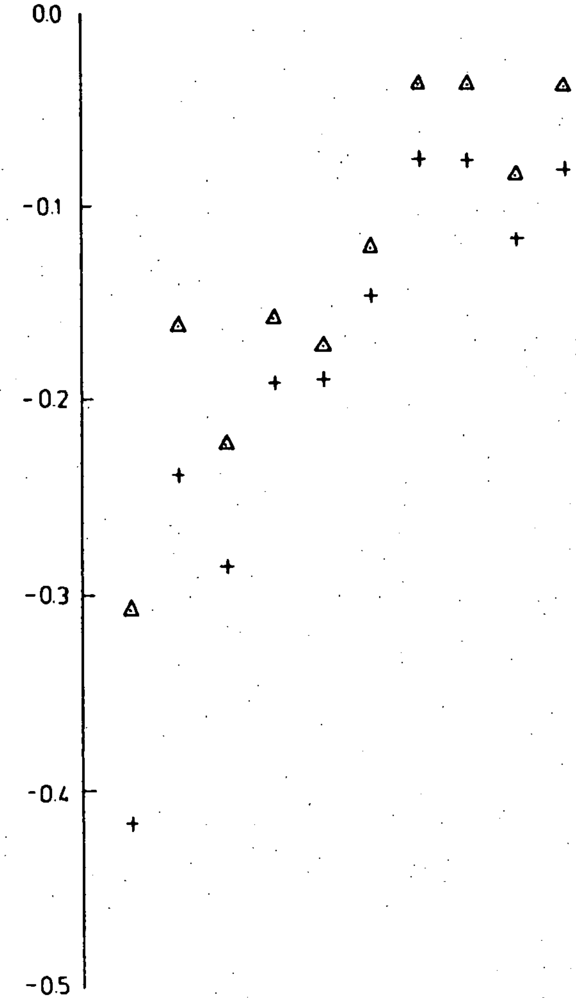


fig. 4.36 $S_3(0.114)$

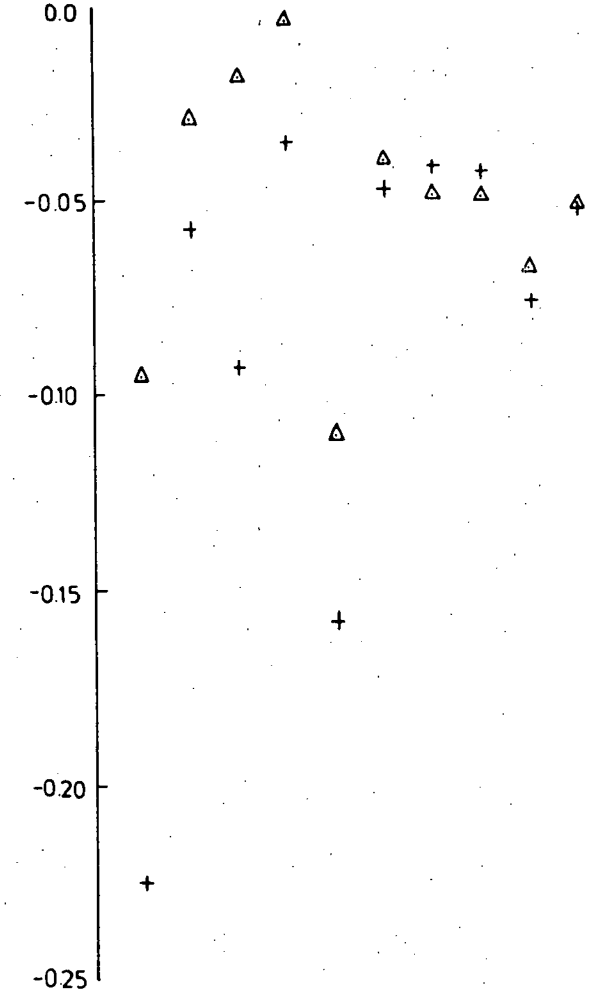


fig. 4.37 $S_3(0.228)$

this particular strategy of automatic design to be of any great value. Its introduction was for the purpose of providing some automatic design process which used overall assessment quantities based on the spot diagram and derivatives of these quantities. The background to the strategy was explained in the introductory section of this chapter.

Consider now the variation of the quantities P_1, Var, $S_1(r)$ and $S_{1M}(r)$ during the course of optimization. Firstly, all assessments indicate a strong improvement in the first three steps and then only marginal changes, indicating broad agreement between different assessment functions. However the correction process over stages four, five and six shows some rather surprising results. The value of P_1 for each of the field angles shows a marked increase at stage five which is not reflected in the other assessment quantities except, to a very slight extent, in the axial values for S_1 and S_{1M} at lowest spatial frequency (r=0.057). In fact all other asesssments indicate a marked improvement in the state of correction. As mentioned in the previous chapter the values of S_1 derived from o.t.f. values less than about 0.2 are of little value, hence S_1 (0.228) should be ignored for the 7.5° values. It can also be seen that S_{1M} (0.114, S_1 (0.114), S_{1M} (0.228) and Var all show agreement regarding the axial correction states, and S_1 (0.228) substantially agrees with these, apart from the initial values. However an examination of optimum image plane positions for the systems arising at the various stages of design shows that much of the ranking using S_1 (0.228) is determined by the distance of the optimum image plane from the plane at which evaluation is taking place. For lower spatial frequencies the values of S_1 are not so rapidly varying with image plane position, as indicated by fig. 3.1. Use of the variance, $V(r,\psi)$, of the difference function thus enables one to obtain results corresponding to use of P_1 or Var or intermediate functions. Furthermore the quantity $V(r,\psi)$ may be readily evaluated at the optimum plane since this plane is quickly calculated using the expression presented in the previous chapter. We see from this analysis that use of P_1 as a function for optimization in automatic design programs provides results which are in broad agreement with many other assessment quantities, though it would appear that some differences would arise in the latter stages of correction if one were to use criteria based on Var or spatial frequency functions evaluated near the limit of the useful bandwidth. These problems may be eliminated by use of the quantity S_{1M} modified to make allowance for products of two negative quantities.

Turning now to the astigmatism quantities P_2, S_2 and S_{2M} we find that P_2 decreases towards zero as also do S_2 (0.057) and S_{2M} (0.057) though the approach to zero differs somewhat between P_2 and the other functions. When S_2 and S_{2M} are evaluated at higher spatial frequencies the results are not as satisfying, indicating that this quantity may not be a very useful function for astigmatism determination at higher spatial frequencies.

The plots of S_3 indicate the systematic improvement in asymmetry during the design process. It can be seen that S_3 (0.057) and S_3 (0.114) are in agreement while S_3 (0.228) differs quite considerably. This would suggest that S_3, like S_2, ought to be calculated at low spatial frequencies.

b) <u>Optimization using Coefficient Reduction and Balancing</u>

The graphs of values of the various assessment functions at successive stages of this optimization are shown in figs. 4.38 to 4.55. The system was evaluated in the gaussian image plane in mercury green light (5461Å) and for a focal length of 24 inches. All quantities are calculated in units of the focal length. Results are given for angles of 0°, 4°, 7° and 10° off-axis.

Consider firstly those functions measuring image spread - namely P_1, Var, S_1 and S_{1M}. The design procedure is based, ultimately, on reducing the spread in transverse aberrations as indicated by the spot diagram. From the plot of P_1 values it would appear that much of the design process

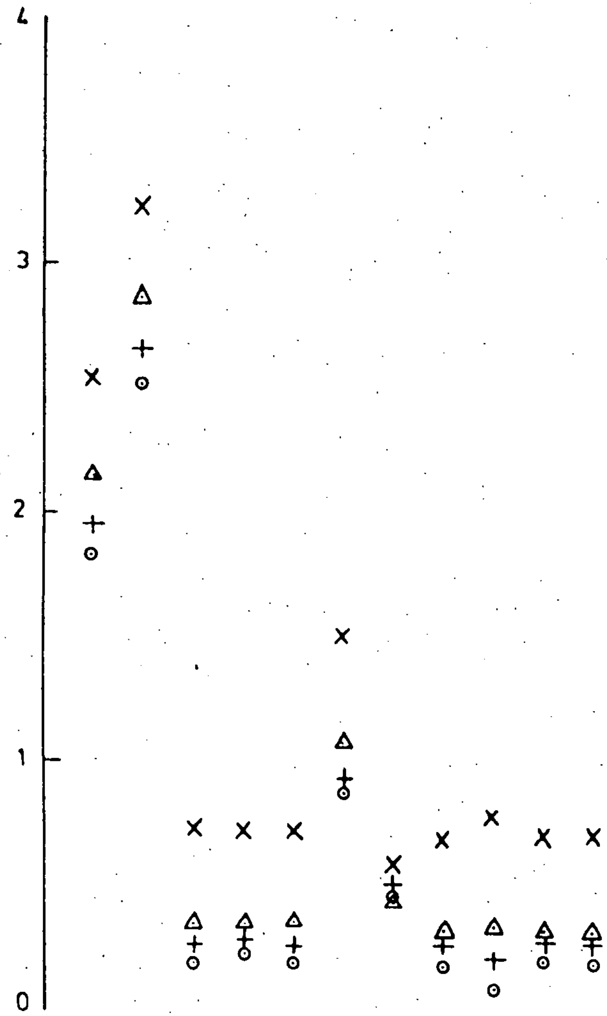


fig. 4.38 $P_1 \times 10^4$

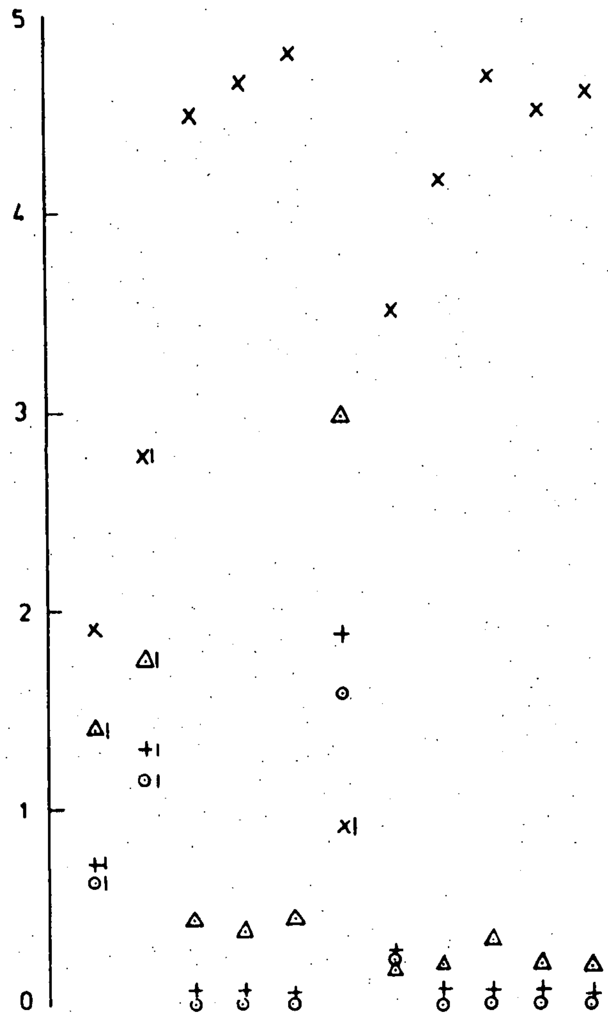


fig. 4.39 Wavefront Variance

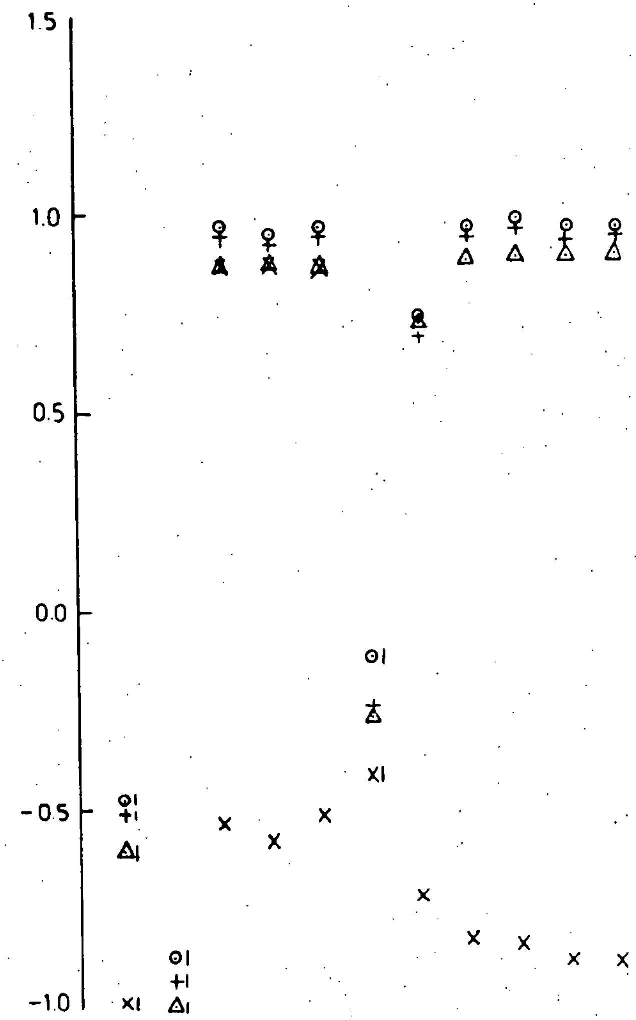


fig. 4.40 $S_{1M}(0.022)$

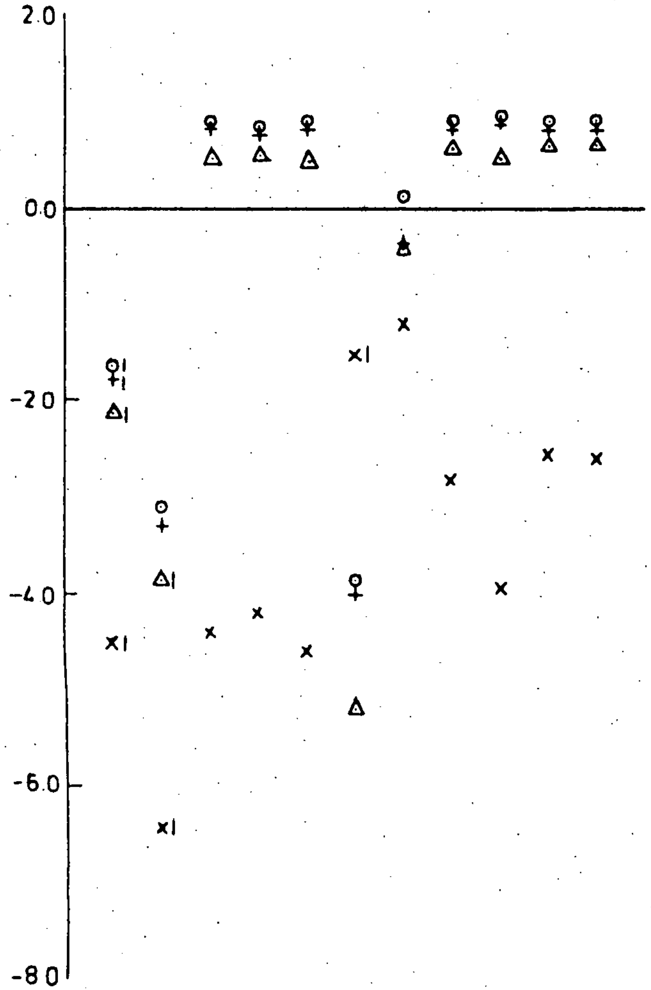


fig 4.41 $S_{1M}(0.044)$

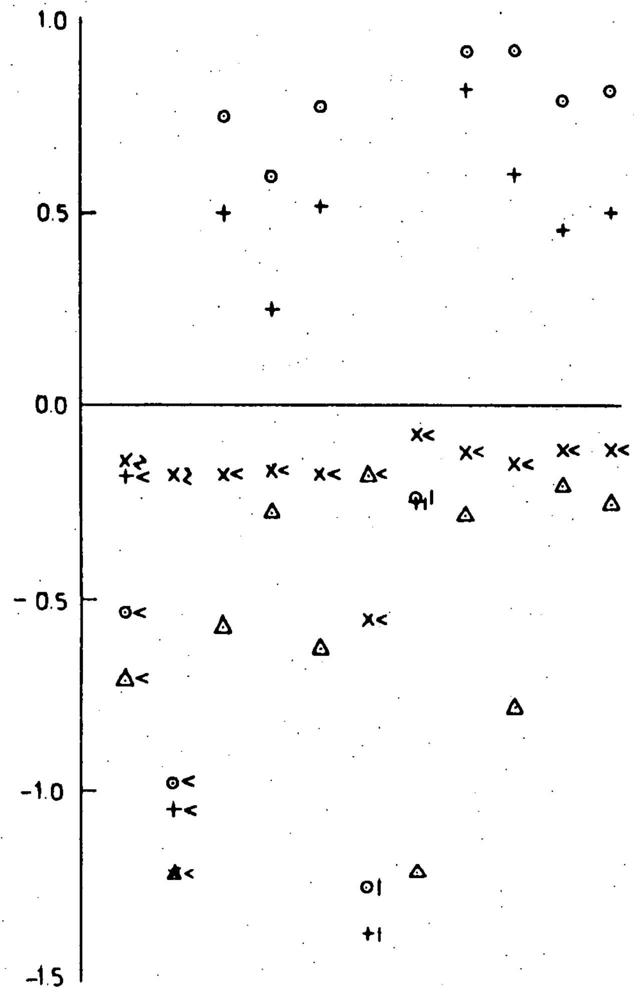


fig. 4.42 $S_{1M}(0.087)$

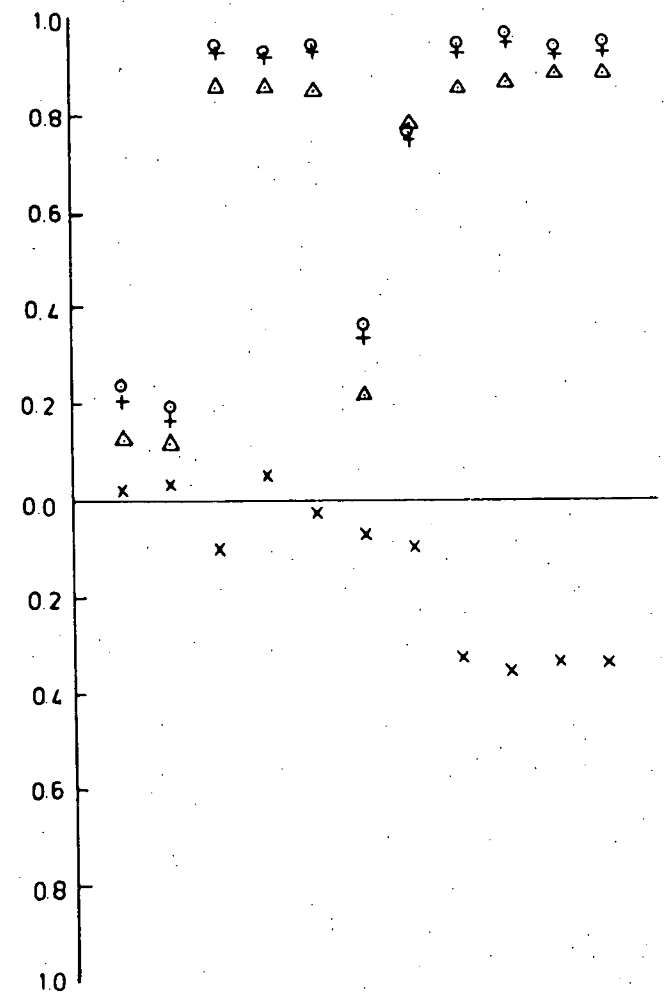


fig 4.43 $S_1(0.022)$

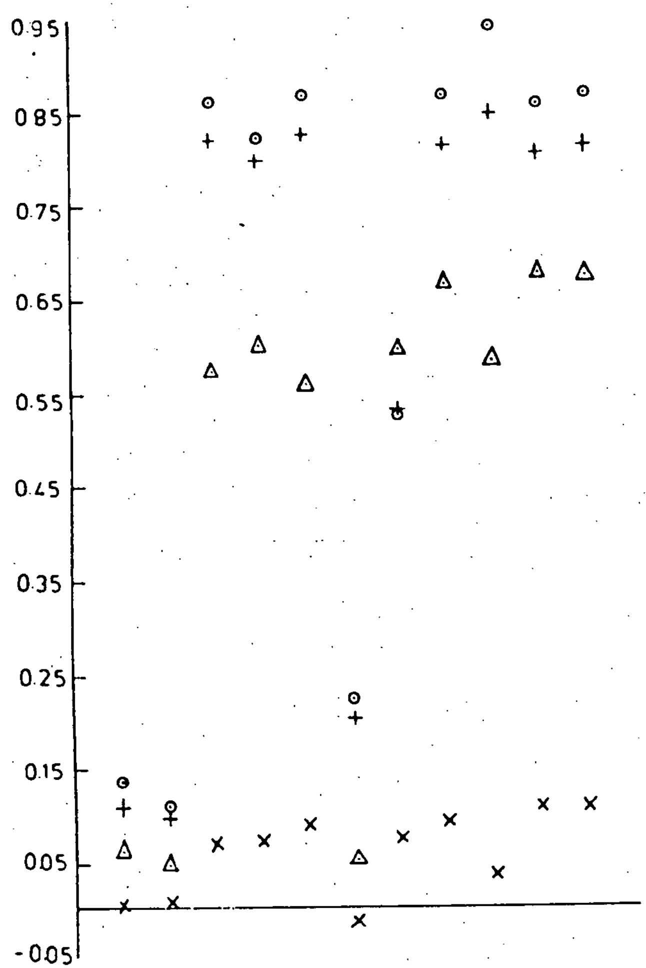


fig. 4.44 $S_1(0.044)$

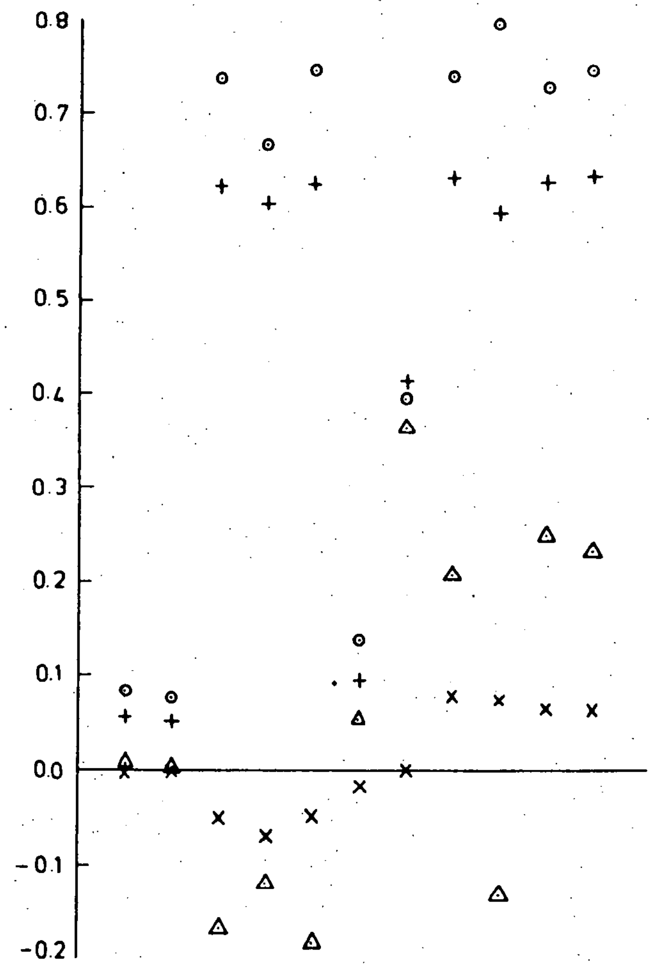


fig. 4.45 $S_1(0.087)$

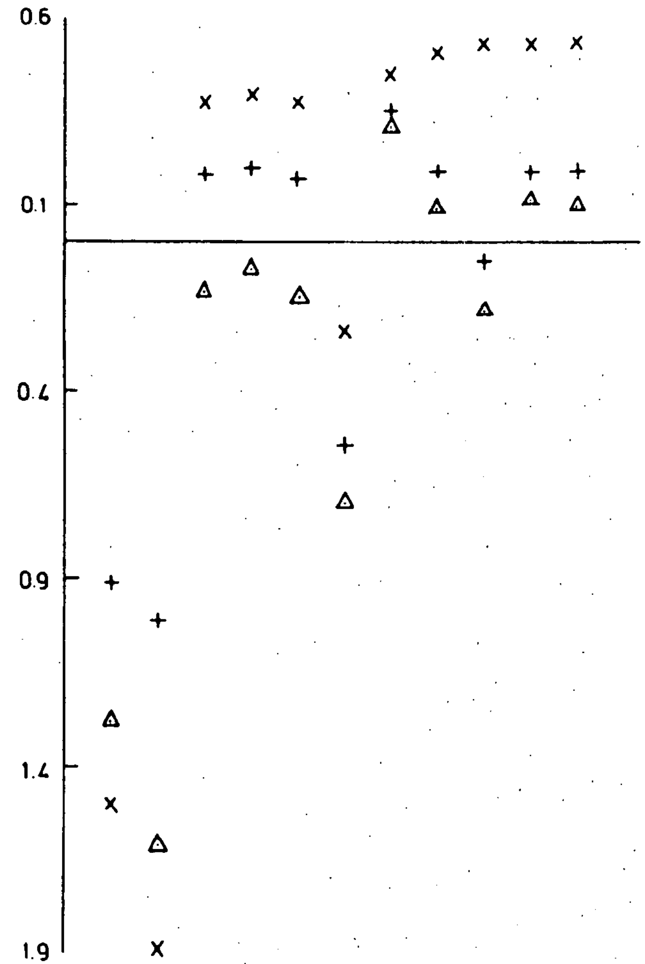


fig. 4.46 $P_2 \times 10^4$

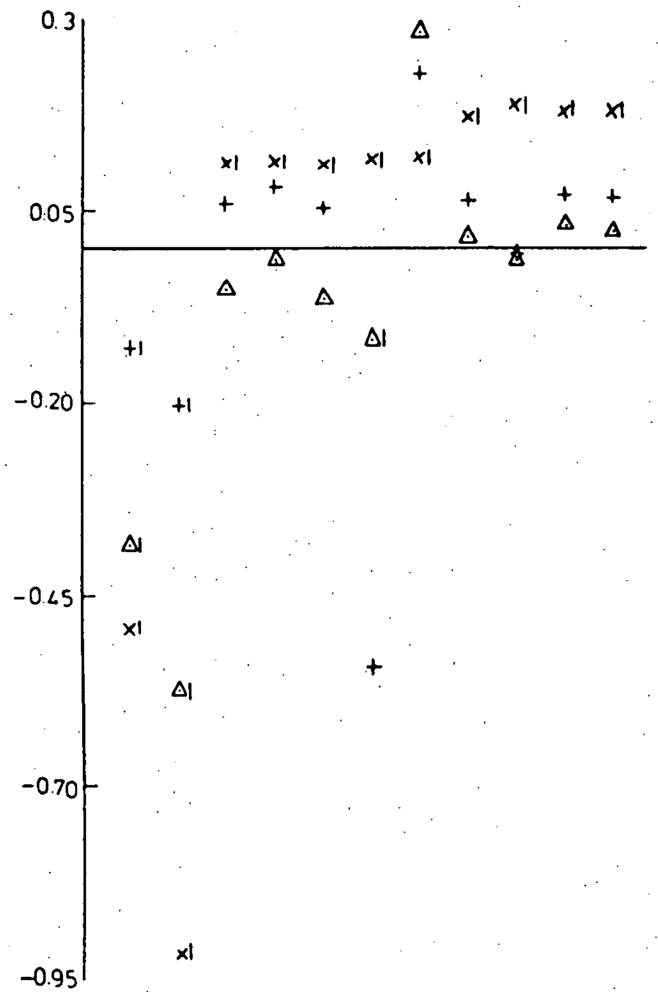


fig. 4.47 S_{2M} (0.022)

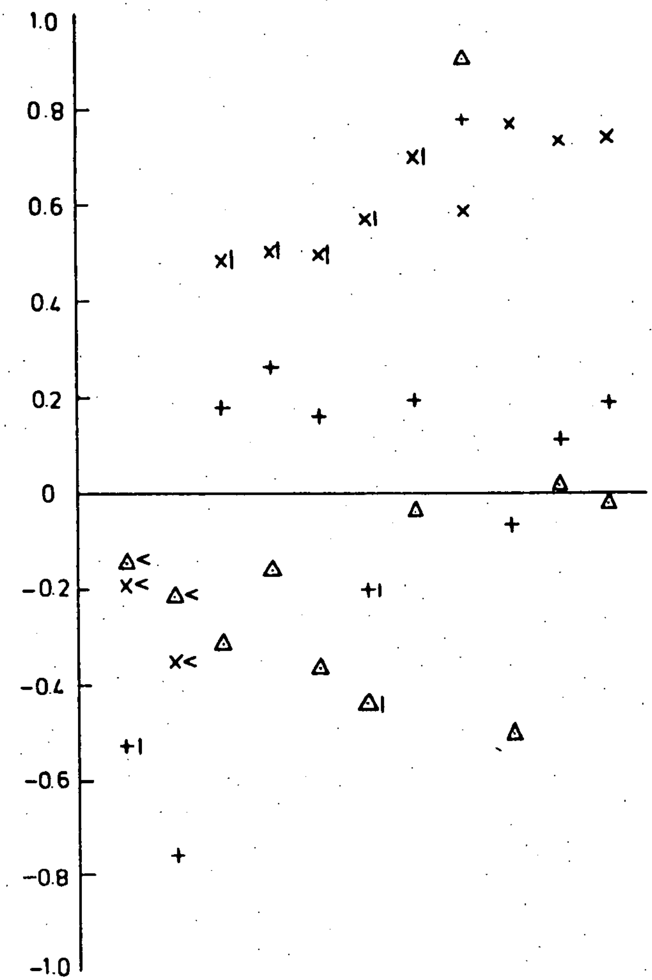


fig. 4.48 S_{2M} (0.044)

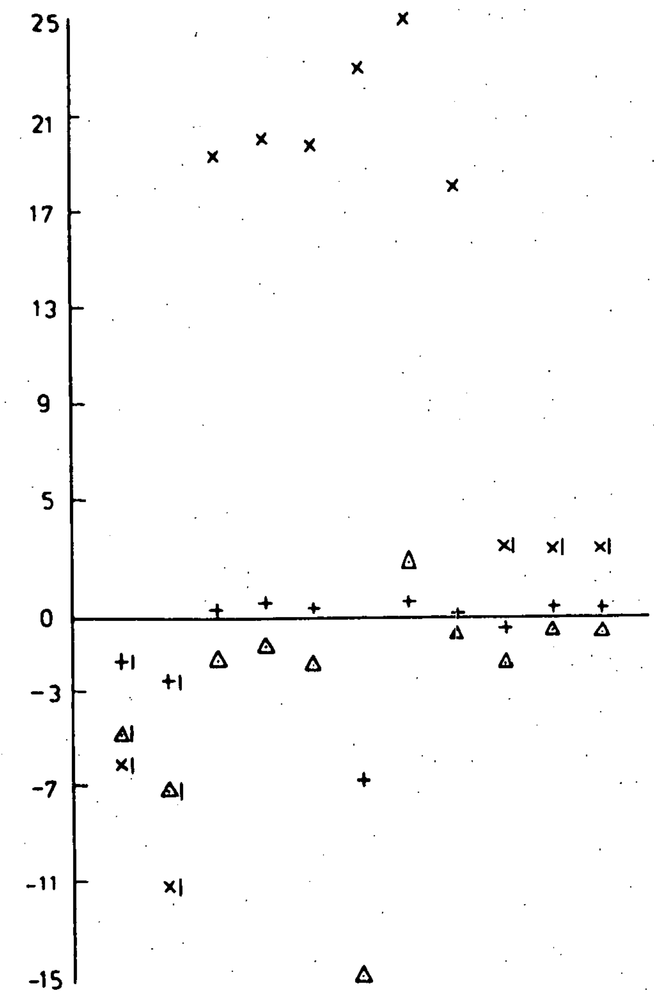


fig. 4.49 S_{2M} (0.087)

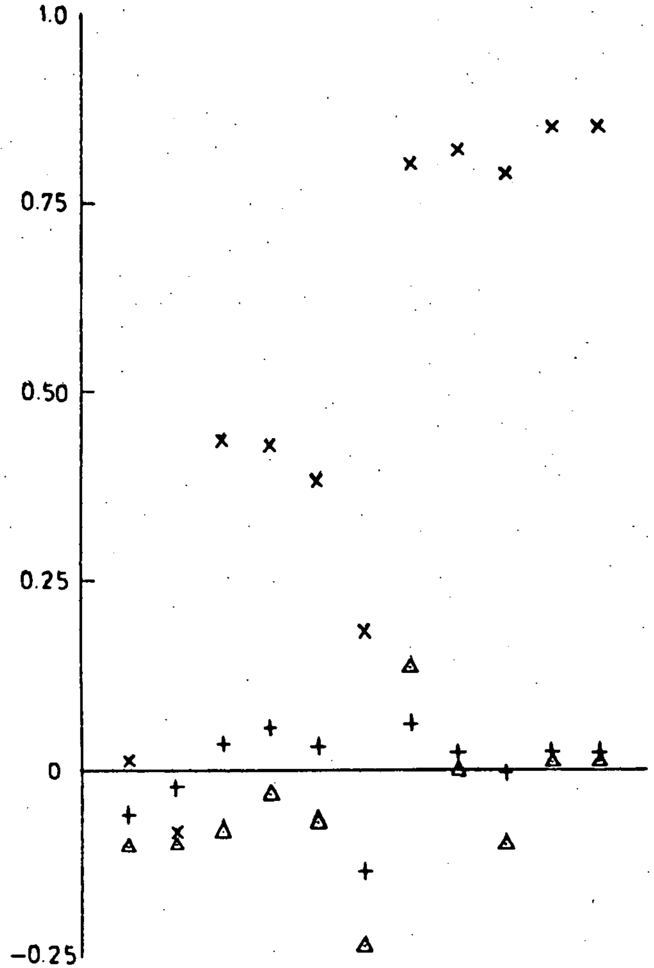


fig. 4.50 $S_2(0.022)$

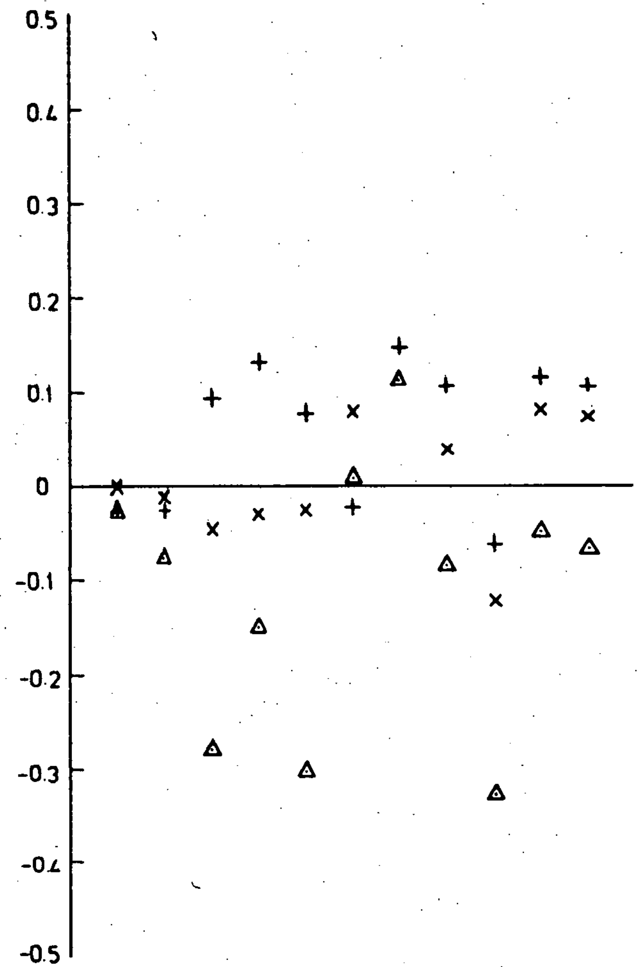


fig. 4.51 $S_2(0.044)$

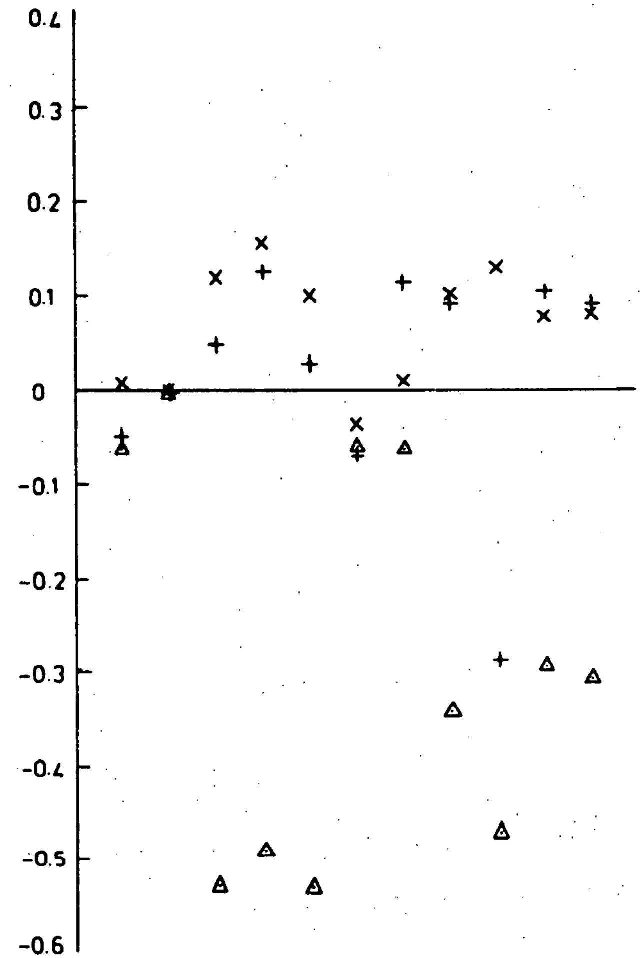


fig. 4.52 $S_2(0.087)$

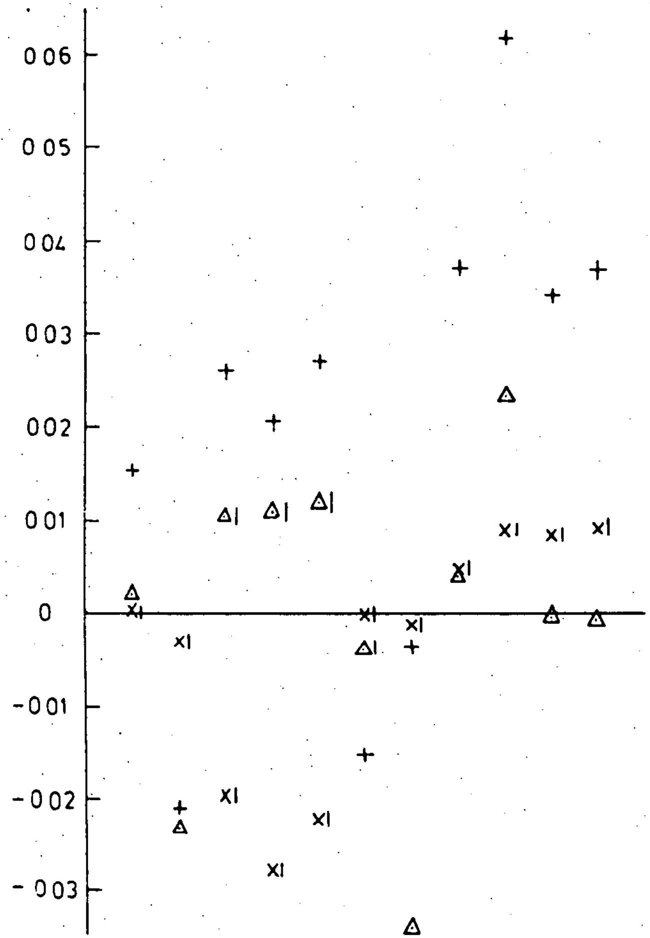


fig. 4.53 S_3(0 022)

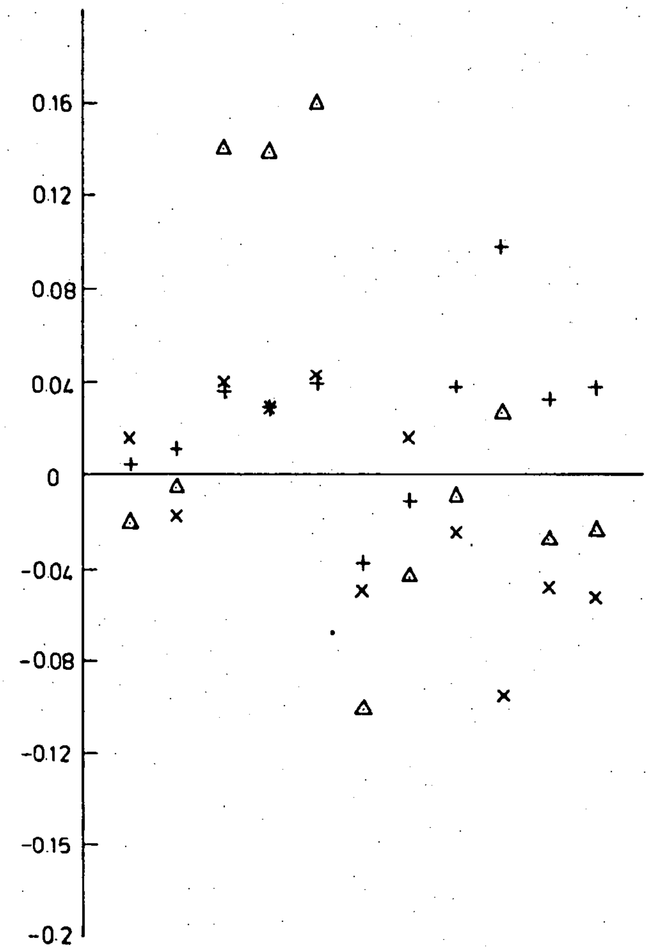


fig. 4.54 S_3(0.044)

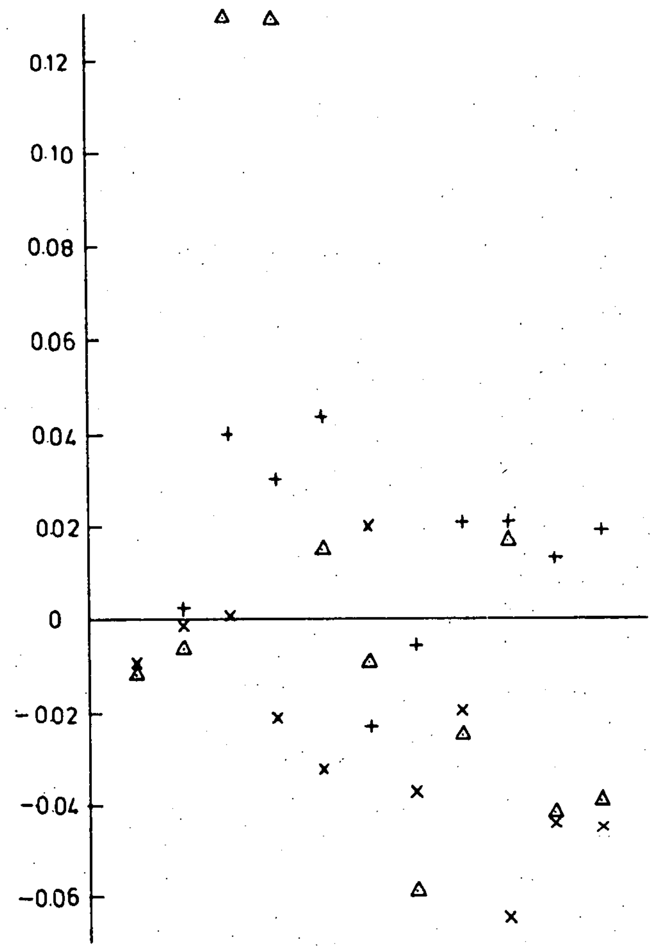


fig. 4.55 S_3(0.087)

is wasted time since there is practically no difference in P_1 between stages 3 and 11, and, apart from stages 6, 7 and 9, what changes were tried made no significant difference. There is, however, a slight improvement shown between stages 3 and 11. These results are substantially supported by the plot of P_2 values, though the 7.5° value of P_2 changes sign between stages 3 and 11. With respect to P_2 it is worth noting that stage 3 actually provides a more satisfactory correction state than stage 11. These results are reflected in the functions Var, S_1 and S_{1M} with some of the differences between stages 3 and 11 becoming more marked in S_1 and S_{1M} as the frequency is increased.

The functions S_2 and S_{2M} show similar behaviour with respect to P_2 as was found for previous assessments - general agreement with P_2 for low spatial frequencies, but results of little value when the o.t.f. falls below about 0.5. A far more useful astigmatism function would be that defined in terms of the distance between sagittal and tangential focii. This has been mentioned earlier, in the comparison of assessment functions.

§4.5 Conclusions

The results presented in this chapter indicate that most assessment functions show general agreement regarding the ranking of states of correction of optical systems arising during the intermediate stages of automatic design. Hence the choice of assessment criterion is not critical in determining optimum <u>regions</u> of parameter space. However details of ranking do differ, and hence true optima can only be defined in terms of a specific assessment function. Furthermore merit functions formed by a weighted summing of values across the field should be treated with considerable caution. This is because the relative values assigned by different assessment functions to any two correction states are, in general, quite different even though the ordering of a given set of correction states is in agreement. This means that a weighted sum of values across the field using one criterion will give a different - in some cases much different - effective field weighting from another function.

Examination of the triplet optimization using Cruickshank's method has suggested that much of the computing done provides only marginal improvement in the overall state of correction. This may, perhaps, be largely overcome by relaxing tolerances on the various iterations, though then the usefulness of some stages - such as modifying the residuals R_2 and R_5 - may be queried.

Finally the potential of the quantity S_{1M}, determined using the variance of the aberration difference function, must be re-emphasized. Rapid computation, both of the function itself and of the optimum image planes using minimization of this variance, together with its versatility make it a most promising assessment parameter for automatic design.

References

4.1) H.A. Buchdahl, Optical Aberration Coefficients, (Dover Publications Inc., New York, 1968).

4.2) F.D. Cruickshank, private communication.

4.3) P.J. Sands, Appl. Opt., 9, 828 (1970).

4.4) B. Brixner, Appl. Opt., 5, 1948 (1966); 12, 2703 (1973).

4.5) M.J. Kidger and C.G. Wynne, Opt. Acta, 14, 279 (1967).

4.6) N.N. Vilchevskaya and N.V. Tseno, Opt. Spectrosc., 33, 179 (1972).

4.7) W.B. King, J.O.S.A., 59, 1155 (1969).

4.8) H. Sumita, Japan J. Appl. Phys., 8, 1027 (1969).

4.9) Y. Itoh, J.O.S.A., 61, 302 (1971).

4.10) D.S. Grey, Appl. Opt., 9, 523 (1970).

4.11) M. Rimmer, Appl. Opt., 9, 533 (1970).

4.12) H. Josefsen, B.I.T., 3, 189 (1964).

4.13) H.D. Taylor, Trans. of Opt. Society, 22, 143 (1923).

4.14) R.E. Hopkins, Japan J. Appl. Phys., 4, Suppl. 1, 60 (1965); also J.O.S.A. 52, 172 (1962).

4.15) A.L.H. Aldersey, Thesis, University of Tasmania, 1967.

GLOSSARY

Through the text there occur many different symbols. To facilitate the reading of this thesis, therefore, a glossary of the more important symbols is included here. Those symbols which occur in only one section are generally not included.

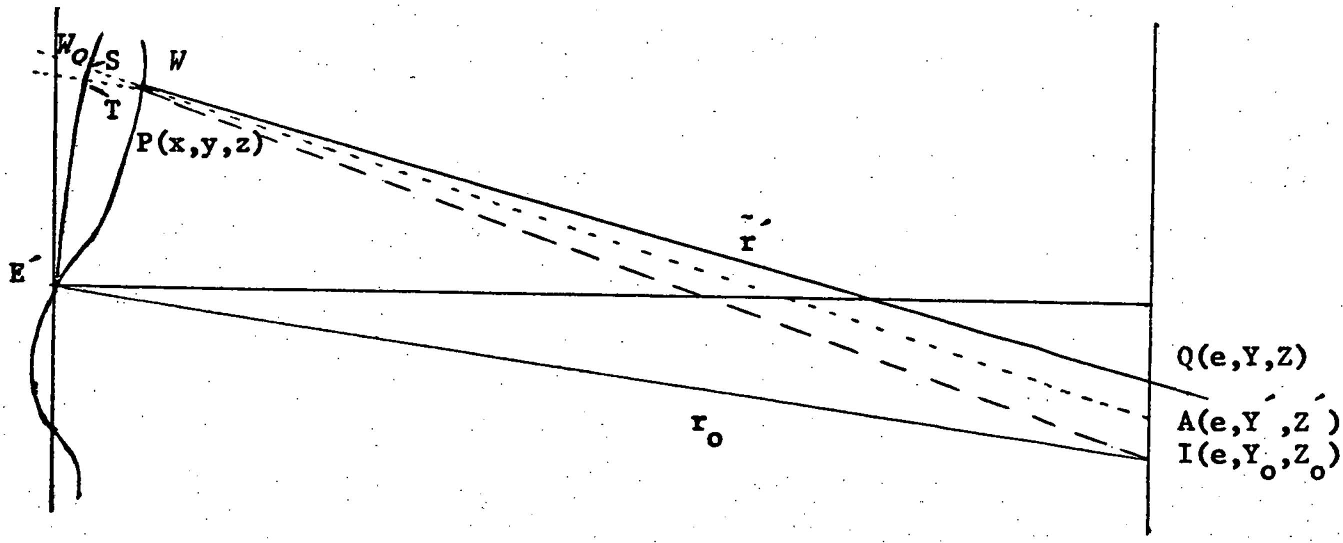
$A, \bar{A}$	third order aberration coefficients;
$A(s,t)$	area common to two replicas of the pupil having centres at $(\mp s, \mp t)$;
$B, \bar{B}$	third order aberration coefficients;
$C, \bar{C}$	" " " " ;
$c_{oj}(j=1, \ldots k)$	axial curvature of the j^{th} surface;
$d_j(j=1, \ldots k)$	axial separation of surfaces $(j-1)$ and j;
$D(\lambda,\mu,\nu)$	deformation function;
e	axial distance between paraxial exit pupil and gaussian image planes;
e_{ij}	$\rho^i \cos^j\theta$;
E_{ij}	$\int_0^{2\pi}\int_0^1 e_{ij}\rho d\rho d\theta$;
$F_k(k=1,\ldots 6)$	spatial frequency functions
$G^{(n)}_{\mu\nu k}$	coefficients in expansion of $\underline{\hat{\varepsilon}}$
$\underline{h}$	paraxial image height (always used with $h_z=0$);
$\underline{H}'_j$	image height $(=\ell_{oj}\underline{V}'_j-\underline{Y}'_j)$;
$\underline{I}_j$	$c_{oj}\underline{Y}_j-\underline{V}_j$;
k	number of surfaces in an optical system;
k	$= 2\pi/\lambda$;
ℓ_{oj}	y_{oj}/v_{oj};
m	paraxial magnification of a system for a given object plane;

m_p	paraxial magnification of a system for pupil planes;
m^*	$= m\ell_{01}$;
N_j	refractive index of medium preceeding surface j;
$\bar{N}_j$	$= N_{j+1} - N_j$;
p	axial distance from front surface to paraxial entrance pupil;
p'	axial distance from back surface to paraxial exit pupil;
$p_i (i=1,...5)$	third order effective aberration coefficients;
$P_j (j=1,2,3)$	spot diagram assessment functions;
$R(\lambda,\mu,\nu)$	wavefront retardation function;
r	fractional spatial frequency;
$\underline{S}$	paracanonical coordinates, first pair;
$s_i (i=1,...12)$	fifth order effective aberration coefficients;
s	spatial frequency for sagittal orientations of elements;
$S_i (i=1,...12)$	fifth order aberration coefficients;
$S_j (j=1,2,3)$	assessment functions based on o.t.f. values;
$\underline{T}$	paracanonical coordinates, second pair;
$t_i (i=1,...20)$	seventh order effective aberration coefficients;
t	spatial frequency for tangential orientations of elements;
$T_i (i=1,...20)$	seventh order aberration coefficients;
$T(r,\psi)$	optical transfer function (o.t.f.);
$\underline{V}_j$	direction tangents of a ray incident on surface j;
V	$= \underline{V}_1$;
$V(r,\psi)$	variance of the aberration difference function;
$W(\lambda,\mu,\nu)$	$= R(\lambda,\mu,\nu)$;
x'	axial distance of image plane from gaussian image plane;

$\underline{Y}_j$	y(and z)-coordinate of point of intersection with j^{th} polar tangent plane;
$z_i(i=1,2)$	defocussing coefficients (longitudinal and transverse);
α	paraxial exit pupil radius
$\Delta(s,t)$	wavefront aberration difference function, $W(y+s,z+t)-W(y-s,z-t)$;
$\underline{\varepsilon}'$	aberration of a ray;
ψ	angle of orientation of line elements and an object, where lines lying parallel to the plane defined by the object point and principal axis have $\psi=0$;
λ	wavelength;
$\underline{\Lambda}$	quasi-invariant, $N(y_o\underline{V}-v_o\underline{Y})$;
μ	yh
$\mu_j(j=1,\ldots 12)$	s_j;
ν	h^2
$\pi_j(j=1,\ldots 5)$	fourth order retardation coefficients;
ρ	radial polar coordinate in first polar tangent plane;
θ	angular polar coordinate;
$\sigma_j(j=1,\ldots 5)$	p_j;
$\sigma_j(j=1,\ldots 9)$	sixth order retardation coefficients;
$\underline{\sigma}$	semi-axes of ellipse used to approximate sheared pupil region;
τ	generic symbol for parameters of the system;
$\tau_j(j=1,\ldots 20)$	t_j;
$\tau_j(j=1,\ldots 14)$	eighth order retardation coefficients;
ξ	$S_y^2 + S_z^2$;
χ	$\sigma_y^2 - \sigma_z^2$;
η	$S_yT_y + S_zT_z$;
ζ	$T_y^2 + T_z^2$.

APPENDIX I. A Discussion of Definitions of the Wavefront Aberration and their Application in Evaluating the Scalar Diffraction Integral.

Consider the following two definitions of the wavefront aberrations using fig. 1:-



TPQ is a ray from an object point O. I is the image point corresponding to the object at O. SPI is a radius of a reference sphere W_o and W is the aberrated wavefront from a point object at O. A is any point in the neighbourhood of I.

The first definition is that given by Rayces,[1] following Nijboer, and is also a more general form (in its choice of reference sphere) of that given by Buchdahl in OAC VII[2]. This is that the wavefront aberration is

$$W^*(y,z) = SP \tag{1}$$

The second definition is that used in this thesis and given by Born and Wolf[3], among others. This gives the aberration as

$$W(y,z) = TP \tag{2}$$

Buchdahl has shown that, when I coincides with the gaussian image point, the two definitions differ by terms of order ten in the co-ordinates (y,z,h) where h is the gaussian image height.

We are not directly concerned here with differences in computation which might arise from using one or the other definition, but rather which definition may be considered the more "correct" when considered from the point of view of its use in optical theory.

The wavefront aberration is introduced because of its usefulness in the scalar theory of diffraction images. Suppose the amplitude of the disturbance at IA due to a source at 0 is required. Then the Fresnel-Kirchhoff formulation of the diffraction integral allows us to write

$$U(\underline{Y}') = \frac{i}{\lambda} \iint_{\text{aperture}} U(\tilde{\underline{y}};\underline{Y}').\ h(\underline{y};\underline{y_o})\ dydz \tag{3}$$

where $U(\underline{y};\underline{Y}')$ gives the contribution to the amplitude at $A(e,\underline{Y}')$ due to a unit amplitude point source at $P(x,y,z)$, and $h(y,z;y_o,z_o)$ describes the amplitude of the disturbance at P due to the source at $O(l_o,y_o,z_o)$. If the aperture illumination is uniform then h is a constant for a given source. This will be assumed in the following. Let $r_o = E'I$, $\tilde{r} = PI$, $\tilde{r}' = PA$ and $X = \angle APQ$. The Huygens-Fresnel principle gives

$$\tilde{U}\ (\underline{y};\ \underline{Y}') = \exp(jk\tilde{r}').\ \frac{\cos X}{\tilde{r}'}\ , \tag{4}$$

where $k = 2\pi/\lambda$, so that

$$U\ (\underline{Y}') = C.\ \iint_{\text{aperture}} \exp(j\ k\ \tilde{r}')\ \frac{\cos X}{\tilde{r}'}\ dydz \tag{5}$$

Now let $\underline{\delta} = \underline{Y}' - \underline{Y}_o$. Then $\cos X = 1 - O\left((\delta/r_o)^2\right)$, and $\tilde{r}' = r_o - O\left(W(y,z)\right)$, where $W(y,z)$ is the wavefront aberration (under either definition). The term $\cos X/\tilde{r}'$ will therefore be a very slowly varying quantity in optical instruments and is taken to be constant. Hence (5) may be written as

$$U(\underline{Y}') = \frac{C \exp(jkr_o)}{r_o} \iint_{\text{aperture}} \exp\left[jk(\tilde{r}' - r_o)\right] dydz \tag{6}$$

Consider now the definitions (1) and (2) in the light of (6). If the point A coincides with I then Rayces' definition (1) gives the quantity $r' - r_o (=\Delta r)$ directly. Using definition (2) however requires that we first assume that the relative phase associated with points on W_o is determined by their distance from W along the direction of energy flow. Under this assumption it is possible to determine the complex amplitude at any point A in the neighbourhood of I.

Consider the calculation of Δr.

$$\tilde{r}'^2 = (e-x)^2 + (y - Y')^2 + (z-Z')^2, \quad r_o^2 = e^2 + Y_o^2 + Z_o^2.$$

The deformation function was introduced in § 1.3 and is defined by the equation

$$(e-x)^2 = e^2 - (y^2+z^2) + 2(y\,Y_o + z\,Z_o) - 2eD$$

Hence $$\tilde{r}'^2 = e^2 + Y_o^2 + Z_o^2 - 2[(y-Y_o).\,\delta_y + (z-Z_o)\delta_z] + \delta_y^2 + \delta_z^2 - 2eD \tag{7}$$

Let $a = y\delta_y + z\delta_z$, $b = Y_o\,\delta_y + Z_o\,\delta_z$, $c = \delta_y^{\,2} + \delta_z^{\,2}$. Then

$$\tilde{r}' - r_o = r_o\left[1 - \frac{2(a-b)}{r_o^2} + \frac{c}{r_o^2} - \frac{2eD}{r_o^2}\right]^{\frac{1}{2}} - r_o$$

$$\text{i.e.}\quad \Delta r = -\frac{a-b}{r_o} - \frac{eD}{r_o} + \tfrac{1}{2}\frac{e^2D^2}{r_o^3} + \frac{e(a-b)D}{r_o^3} + \frac{D}{r_o} + \tfrac{1}{2}\frac{(a-b)^2}{r_o^3} + O(9).$$

Now the wavefront aberration W_B is related to the deformation D by

$$2eD = (2r_o - W_B).\, W_B$$

$$\text{Hence}\quad \Delta r = -\frac{1}{r_o}\left\{(a-b) + W_B\,(r_o + \tfrac{1}{2}W_B - \frac{(a-b)}{r_o^3} - \frac{C}{r_o} - \frac{(a-b)^2}{2r_o^3})\right] + O(9). \tag{8}$$

Using this expression for Δr we can rewrite (6) as

$$U(\underline{Y}') = C.\iint_{\text{apert}} \exp\left\{-jk\,W_B[1+\tfrac{1}{2}\frac{W_B}{r_o} - \frac{(a-b)}{r_o^2} - \frac{c}{r_o^2} - \frac{(a-b)^2}{2r_o^4} + O(9)\right\}$$

$$.\exp\left\{\frac{-jk}{r_o}(a-b)\right\}dydz \tag{9}$$

By neglecting terms of O(9) and redefining the wavefront aberration function as

$$W = W_B\ [\ 1 + \tfrac{1}{2}\frac{W_B}{r_o} - \frac{(a-b)}{r_o^2} - \frac{c}{r_o^2} - \frac{(a-b)^2}{2r_o^4}\ ] \tag{10}$$

equation (9) reduces to the Fraunhhofer form of the diffraction integral. The usual Fraunhhofer approximation assumes $W = W_B$ rather than the expression given in (10). Since a is of the same order as W in magnitude provided $|\underline{\delta}| \leq |\underline{\varepsilon}|$, the error in using the normal form of the diffraction integral is determined by the magnitude of W. Since diffraction only becomes significant when W is small it is clear that, for investigations of the diffraction image in the neighbourhood of the gaussian image, the approximation

$W = W_B$ is quite adequate.

This analysis then provides justification for writing

$$U(\underline{Y}') = K \iint_{\text{apert}} \exp\left\{-jk\,W\right\} \exp\left\{-\frac{jk}{r_o}(a-b)\right\} dydz \tag{11}$$

thus showing that the definition (2) provides a simple means of calculating the complex amplitude in the neighbourhood of the focus of an optical system.

REFERENCES:

A.1) J.L. Rayces, Opt. Acta, 11, 85 (1964)
A.2) H.A. Buchdahl, J.O.S.A., 50, 539 (1960)
A.3) M. Born and E. Wolf, Principles of Optics (Pergamon Press)

APPENDIX 2. Reduction of Round-off Error in Ford's Raytracing Scheme.

A ray-tracing scheme given by Ford[1] for the calculation of transverse ray aberrations has been modified to allow calculation of the wavefront aberration. The method is to simply trace a principal (or reference) ray through the system, storing each of the distances from surface to surface. Then any other ray is traced and the differences between the intersurface distances of this ray and the reference ray are summed. Projection of the emergent ray onto a reference sphere in the exit pupil allows determination of optical path differences.

It was found that in certain cases fluctuations in the wavefront aberration as large as 0.1λ occurred and detailed checking isolated one equation as being mainly responsible. Referring to Ford's paper §3 following equation (5) there occurs the equation (see also Table II, term t_{23})

$$x = (1 - B)/c \tag{1}$$

$$\text{where } B = [\,(p_1{}^2 - p_o p_2)^{\frac{1}{2}} - p_1]/p_2 \tag{2}$$

$$\text{and } p_2 = 1 + \zeta\,,\ p_1 = c\eta - \zeta,\ p_o = c^2\xi - 2p_1 - p_2 \tag{3}$$

It can be seen that if $c < 1$ and $B \approx 1$ then round=off error in the computation of B will be magnified. This method of calculating x, the distance of the point of intersection of a ray with a surface from the polar tangent plane for that surface, was found to be the cause of most of the round-off error.

The computer for which the program was written had 39 bit words and, working in floating point arithmetic, gave relative round-off of $\sim 2 \times 10^{-9}$. The following method eliminates any increase in the magnitude of the round-off error, and involves a minimal number of additional arithmetic operations.

From (1) and (2)

$$1 - B = \{p_2 - p_1 [(1 - p_o p_2/p_1^2)^{\frac{1}{2}} - 1]\}/p_2$$

Now $\quad p_1^2 - p_o p_2 = p_1^2 - (c^2\xi - 2p_1 - p_2)\, p_2$

$$= (p_1 + p_2)^2 - c^2\xi\, p_2$$

so that $\quad 1 - B = (\frac{p_1 + p_2}{p_2}).\{1 - [1 - \theta/(p_1 + p_2)^2\,]^{\frac{1}{2}}\,\}$

$$\text{where } \theta = c^2\xi\, p_2/(p_1 + p_2)^2 = c^2\xi p_2/(1 + c\eta)^2$$

Using a binomial expansion then gives

$$1 - B = (1 + p_1/p_2).\,\tfrac{1}{2}\theta\,\{1 + 1/4\,\theta + 1/8\,\theta^2 + 5/64\,\theta^3 + \ldots\} \quad (4)$$

Remembering that (1) only becomes unsatisfactory for small c and that ξ, η, and ζ are generally $\lesssim 1$ it is clear that θ, when used, will be small and hence (4) will be rapidly convergent. Hence to make the first neglected term within round-off if this term is the θ^4 term requires that $\theta \leq 1.5 \times 10^{-2}$. In the axial case this requires that (4) be used in the computation of x whenever $c \leq 0.122$.

A2.1) P.W. Ford, J. Opt. Soc. Am., 50, 528. (1960).

APPENDIX 3.

Tertiary Terms for the Equivalent Coordinate Variations.

The equivalent coordinate variations are given by the expression

$$\frac{\partial \hat{\varepsilon}'_y}{\partial x} \cdot \frac{\partial x}{\partial c_{oj}} = \Big[\, \Gamma a\left(S_y \frac{\partial}{\partial \xi} + 2\xi \frac{\partial}{\partial \xi} + \eta \frac{\partial}{\partial \eta}\right) + \bar{\Gamma} a\left(T_y \frac{\partial}{\partial S} + 2\eta \frac{\partial}{\partial \xi} + \zeta \frac{\partial}{\partial \eta}\right)$$

$$+ \Gamma b\left(S_y \frac{\partial}{\partial T} + \xi \frac{\partial}{\partial \eta} + 2\eta \frac{\partial}{\partial \zeta}\right) + \bar{\Gamma} b\left(T_y \frac{\partial}{\partial T} + \eta \frac{\partial}{\partial \eta} + 2\zeta \frac{\partial}{\partial \zeta}\right)\Big]\, \hat{\varepsilon}'_y$$

- see equation (18)

Replacing Γa, $\bar{\Gamma} a$, Γb, $\bar{\Gamma} b$ and $\hat{\varepsilon}'$ by their expansions then gives the required expansion. The coefficients of the tertiary terms occuring in this expansion are given below:-

$\underline{\underline{S}}\ \xi_1^3$: $\theta_y . 7\bar{T}_1 + \theta_v(\bar{T}_1 + T_2) + \alpha_y.5S_1 + \alpha_v(\bar{S}_1 + S_2) + \sigma_{1y}.3A + \sigma_{1v}(\bar{A} + B)$.

$\underline{\underline{T}}\ \xi_1^3$: $\theta_y\ 6\bar{T}_1 + \bar{\theta}_y\ T_1 + \theta_v\ \bar{T}_2 + \bar{\theta}_v\bar{T}_1 + 4\alpha_y\bar{S}_1 + \bar{\alpha}_y S_1 + \alpha_v\bar{S}_2 + \bar{\alpha}_v\bar{S}_1$

$+ 2\sigma_{1y}\ \bar{A} + \bar{\sigma}_{1y}\ A + \sigma_{1v}\ \bar{B} + \bar{\sigma}_{1v}\ \bar{A}$.

$\underline{\underline{S}}\ \xi_1^2\eta_1$: $\theta_y.6T_2 + \bar{\theta}_y 6T_1 + \theta_v(\bar{T}_2 + 2T_4 + 2T_3) + \bar{\theta}_v\ T_2 + \alpha_y 4S_2 + \bar{\alpha}_y 4S_1$

$+ \alpha_v(\bar{S}_2 + 2S_4 + 2S_3) + \bar{\alpha}_v S_2 + \sigma_{1y}.2B + \bar{\sigma}_{1y}2A + \sigma_{1v}(\bar{B} + 2C)$

$+ \bar{\sigma}_{1v}B + \sigma_{2y}3A + \sigma_{2v}\ (\bar{A} + B)$.

$\underline{\underline{T}}\ \xi_1^2\eta_1$: $\theta_y\ 5\bar{T}_2 + \bar{\theta}_y\ (T_2 + 6\bar{T}_1) + \theta_v\ (2\bar{T}_4 + 2\bar{T}_3) + \bar{\theta}_v.\ 2\bar{T}_2 + \alpha_y 3\bar{S}_2$

$+ \bar{\alpha}_y\ (S_2 + 4\bar{S}_1) + \alpha_v.2(\bar{S}_3 + \bar{S}_4) + \bar{\alpha}_v\ 2\bar{S}_2 + \beta_y\ 4\bar{S}_1 + \bar{\beta}_y\ S_1$

$+ \beta_v\ \bar{S}_2 + \bar{\beta}_v\ \bar{S}_1 + \sigma_{1y}\ \bar{B} + \bar{\sigma}_{1y}\ (B + 2\bar{A}) + \sigma_{1v}2\bar{C} + \bar{\sigma}_{1v}\ 2\bar{B}$

$+ \sigma_{2y}\ 2\bar{A} + \bar{\sigma}_{2y}\ A + \sigma_{2v}\ \bar{B} + \bar{\sigma}_{2v}\ \bar{A}$.

$$\underline{\underline{S}}\ \xi_1^2\zeta:\ \theta_y\, 5T_3 + \bar{\theta}_y T_2 + \theta_v(\bar{T}_3 + \bar{T}_5) + \bar{\theta}_v T_3 + \alpha_y 3S_3 + \bar{\alpha}_y\, S_2$$

$$+ \alpha_v(\bar{S}_3 + S_5) + \bar{\alpha}_v 2S_3 + \gamma_y 5S_1 + \gamma_v\, (\bar{S}_1 + S_2) + \sigma_{1y}\, C + \bar{\sigma}_{1y} B$$

$$+ \sigma_{1v}\bar{C} + \bar{\sigma}_{1v}\, 2C + \sigma_{3y}\, 3A + \sigma_{3v}\, (\bar{A} + B).$$

$$\underline{\underline{T}}\ \xi_1^2:\ \theta_y 4\bar{T}_3 + \bar{\theta}_y\, (T_3 + \bar{T}_2) + \theta_v \bar{T}_5 + \bar{\theta}_v\, 3\bar{T}_3 + \alpha_y .2\bar{S}_3 + \bar{\alpha}_y\, (S_3 + \bar{S}_2)$$

$$+ \alpha_v\, \bar{S}_5 + \bar{\alpha}_v 3\bar{S}_3 + \gamma_y\, 4\bar{S}_1 + \bar{\gamma}_y S_1 + \gamma_v \bar{S}_2 + \bar{\gamma}_v \bar{S}_1 + \bar{\sigma}_{1y}(C + \bar{B})$$

$$+ \bar{\sigma}_{1v}\, 3\bar{C} + \sigma_{3y}.\, 2\bar{A} + \bar{\sigma}_{3y} A + \sigma_{3v}\, \bar{B} + \bar{\sigma}_{3v}\, \bar{A}\ .$$

$$\underline{\underline{S}}\ \xi_1\eta_1^2:\ \theta_y\, 5T_4 + \bar{\theta}_y\, 4T_2 + \theta_v(\bar{T}_4 + 3T_7 + 2T_5) + \bar{\theta}_v\, 2T_4 + \alpha_y\, 3S_4$$

$$+ \bar{\alpha}_y .2S_2 + \alpha_v(\bar{S}_4 + 2S_5) + \bar{\alpha}_v .2S_4 + \beta_y .4S_2 + \bar{\beta}_y .4S_1$$

$$+\beta_v\, (\bar{S}_2 + 2S_4 + 2S_3) + \bar{\beta}_v S_2 + \sigma_{2y} .2B + \bar{\sigma}_{2y} .2A + \sigma_{2v}(\bar{B} + 2C)$$

$$+ \bar{\sigma}_{2v} B + \sigma_{4y} .3A + \sigma_{4v}\, (\bar{A} + B).$$

$$\underline{\underline{T}}\ \xi_1\eta_1^2:\ \theta_y .4\bar{T}_4 + \bar{\theta}_y\, (T_4 + 2\bar{T}_2) + \theta_v(3\bar{T}_7 + 2\bar{T}_5) + \bar{\theta}_v .3\bar{T}_4 + \alpha_y 2\bar{S}_4$$

$$+ \bar{\alpha}_y\, (S_4 + 2\bar{S}_2) + \alpha_v .2\bar{S}_5 + \bar{\alpha}_v .3\bar{S}_4 + \beta_y .3\bar{S}_2 + \bar{\beta}_y .(S_2 + 4\bar{S}_1)$$

$$+ \beta_v .2(\bar{S}_3 + \bar{S}_4) + \bar{\beta}_v .2\bar{S}_2 + \sigma_{2y}\bar{B} + \bar{\sigma}_{2y}\, (B + 2\bar{A}) + \sigma_{2v} .2\bar{C} + \bar{\sigma}_{2v} .2\bar{B}$$

$$+ \sigma_{4y} .2\bar{A} + \bar{\sigma}_{4y} .A + \sigma_{4v}\bar{B} + \bar{\sigma}_{4v}\, \bar{A}\ .$$

$$\underline{\underline{S}}\ \xi_1\eta_1\zeta_1:\ \theta_y .4T_5 + \bar{\theta}_y\, (4T_3 + 2T_4) + \theta_v(\bar{T}_5 + 2T_8 + 4T_6) + \bar{\theta}_v .3T_5 + \alpha_y .2S_5$$

$$+ \bar{\alpha}_y .2(S_3 + S_4) + \alpha_v(\bar{S}_5 + 4S_6) + \bar{\alpha}_v .3S_5 + \beta_y .3S_3 + \bar{\beta}_y .S_2$$

$$+ \beta_v(\overline{S}_3 + S_5) + \overline{\beta}_v.2S_3 + \gamma_y.4S_2 + \overline{\gamma}_y.4S_1 + \gamma_v(\overline{S}_2 + 2S_3 + 2S_4)$$

$$+ \overline{\gamma}_v S_2 + \sigma_{2y}C + \overline{\sigma}_{2y}B + \sigma_{2v}\ \overline{C} + \overline{\sigma}_{2v}.2C + \sigma_{3y}.2B + \overline{\sigma}_{3y}.2A$$

$$+ \sigma_{3v}\ (\overline{B} + 2C) + \overline{\sigma}_{3v}B + \sigma_{5y}.3A + \sigma_{5v}(\overline{A} + B)$$

$\underline{\underline{T}}\ \xi_1\eta_1\zeta_1$: $\theta_y 3\overline{T}_5 + \overline{\theta}_y(T_5 + 4\overline{T}_3 + 2\overline{T}_4) + \theta_v(4\overline{T}_6 + 2\overline{T}_8) + \overline{\theta}_v.4\overline{T}_5 + \alpha_y\overline{S}_5$

$$+ \overline{\alpha}_y(S_5 + 2\overline{S}_3 + 2\overline{S}_4) + \alpha_v.4\overline{S}_6 + \overline{\alpha}_v.4\overline{S}_5 + \beta_y.2\overline{S}_3 + \overline{\beta}_y(\overline{S}_2 + S_3)$$

$$+ \beta_v.\overline{S}_5 + \overline{\beta}_v.3\overline{S}_3 + \gamma_y.3\overline{S}_2 + \overline{\gamma}_y(4\overline{S}_1 + S_2) + \gamma_v.2(\overline{S}_3 + \overline{S}_4)$$

$$+ \overline{\gamma}_v.2\overline{S}_2 + \overline{\sigma}_{2y}.(C + \overline{B}) + \overline{\sigma}_{2v}.3\overline{C} + \sigma_{3y}\ \overline{B} + \overline{\sigma}_{3y}(2\overline{A} + B) + \sigma_{3v}.2\overline{C}$$

$$+ \overline{\sigma}_{3v}.2\overline{B} + \sigma_{5y}.2\overline{A} + \overline{\sigma}_{5y}A + \sigma_{5v}.\overline{B} + \overline{\sigma}_{5v}.\overline{A}.$$

$\underline{\underline{S}}\ \xi_1\zeta_1^2$: $\theta_y.3T_6 + \overline{\theta}_y T_5 + \theta_v(\overline{T}_6 + T_9) + \overline{\theta}_v.4T_6 + \alpha_y.S_6 + \overline{\alpha}_y.S_5$

$$+ \alpha_v\ \overline{S}_6 + \overline{\alpha}_v.4S_6 + \gamma_y.3S_3 + \overline{\gamma}_y.S_2 + \gamma_v(\overline{S}_3 + S_5) + \overline{\gamma}_v 2S_3$$

$$+ \sigma_{3y}.C + \overline{\sigma}_{3y}B + \sigma_{3v}.\overline{C} + \overline{\sigma}_{3v}\ .2C + \sigma_{6y}.3A + \sigma_{6v}.(\overline{A} + B).$$

$\underline{\underline{T}}\ \xi_1\zeta_1^2$: $\theta_y\overline{T}_6 + \overline{\theta}_y.(T_6 + \overline{T}_5) + \theta_v\overline{T}_9 + \overline{\theta}_v.5\overline{T}_6 + \overline{\alpha}_y\ (S_6 + \overline{S}_5) + \overline{\alpha}_v.5\overline{S}_6$

$$+ \gamma_y.2\overline{S}_3 + \overline{\gamma}_y(\overline{S}_2 + S_3) + \gamma_v.\overline{S}_5 + \overline{\gamma}_v.3\overline{S}_3 + \overline{\sigma}_{3y}.(C + \overline{B})$$

$$+ \overline{\sigma}_{3v}.3\overline{C} + \sigma_{6y}.2\overline{A} + \overline{\sigma}_{6y}.A + \sigma_{6v}.\overline{B} + \overline{\sigma}_{6v}.\overline{A}$$

$\underline{\underline{S}}\ \eta_1^3$: $\theta_y.4T_7 + \overline{\theta}_y.2T_4 + \theta_v.(\overline{T}_7 + 2T_8) + \overline{\theta}_v.3T_7 + \beta_y.3S_4$

$$+ \overline{\beta}_y.2S_2 + \beta_v.(\overline{S}_4 + 2S_5) + \overline{\beta}_v.2S_4 + \sigma_{4y}.2B + \overline{\sigma}_{4y}.2A + \sigma_{4v}(\overline{B}+2C)$$

$$+ \overline{\sigma}_{4v}.B;$$

$$\underline{\underline{T}}\ \eta_1^3 : \quad \theta_y.3T_7 + \overline{\theta}_y(T_7 + 2\overline{T}_4) + \theta_v.2\overline{T}_8 + \overline{\theta}_v.4\overline{T}_7 + \beta_y.2\overline{S}_4 + \overline{\beta}_y.(2\overline{S}_2 + S_4)$$

$$+ \beta_v.2\overline{S}_5 + \overline{\beta}_v.3\overline{S}_4 + \sigma_{4y}.\overline{B} + \overline{\sigma}_{4y}.(2\overline{A} + B) + \sigma_{4v}.2\overline{C} + \overline{\sigma}_{4v}.2\overline{B}.$$

$$\underline{\underline{S}}\ \eta_1^2\zeta_1 : \quad \theta_y.3T_8 + \overline{\theta}_y.(2T_5 + 3T_7) + \theta_v.(\overline{T}_8 + 4T_9) + \overline{\theta}_v.4T_8 + \beta_y.2S_5$$

$$+ \overline{\beta}_y.2(S_3 + S_4) + \beta_v.(\overline{S}_5 + 4S_6) + \overline{\beta}_v.3S_5 + \gamma_y.3S_4 + \overline{\gamma}_y.2S_2$$

$$+ \gamma_v.(\overline{S}_4 + 2S_5) + \overline{\gamma}_v.2S_4 + \sigma_{4y}.C + \overline{\sigma}_{4y}.B + \sigma_{4v}.\overline{C} + \overline{\sigma}_{4v}.2C$$

$$+ \sigma_{5y}.2B + \overline{\sigma}_{5y}.2A + \sigma_{5v}.(\overline{B} + 2C) + \overline{\sigma}_{5v}.B.$$

$$\underline{\underline{T}}\ \eta_1^2\zeta_1 : \quad \theta_y.2\overline{T}_8 + \overline{\theta}_y.(T_8 + 2\overline{T}_5 + 3\overline{T}_7) + \theta_v.4\overline{T}_9 + \overline{\theta}_v.5\overline{T}_8 + \beta_y.\overline{S}_5$$

$$+ \overline{\beta}_y.(S_5 + 2\overline{S}_3 + 2\overline{S}_4) + \beta_v.4\overline{S}_6 + \overline{\beta}_v.\ 4\overline{S}_5 + \gamma_y.2\overline{S}_4 + \overline{\gamma}_y.(2\overline{S}_2$$

$$+ S_4) + \gamma_v.2\overline{S}_5 + \overline{\gamma}_v.3\overline{S}_4 + \overline{\sigma}_{4y}.(C + \overline{B}) + \overline{\sigma}_{4v}.3\overline{C} + \sigma_{5y}.\overline{B}$$

$$+ \overline{\sigma}_{5y}.(B + 2\overline{A}) + \sigma_{5v}.2\overline{C} + \overline{\sigma}_{5v}.2\overline{B}.$$

$$\underline{\underline{S}}\ \eta_1\zeta_1^2 : \quad \theta_y.2T_9 + \overline{\theta}_y.2(T_6 + T_8) + \theta_v.(\overline{T}_9 + 6T_{10}) + \overline{\theta}_v.5T_9 + \beta_y.S_6 + \overline{\beta}_y.S_5$$

$$+ \beta_v.\overline{S}_6 + \overline{\beta}_v.4S_6 + \gamma_y.2S_5 + \overline{\gamma}_y.2(S_3 + S_4) + \gamma_v.(\overline{S}_5 + 4S_6)$$

$$+ \overline{\gamma}_v.3S_5 + \sigma_{5y}.C + \overline{\sigma}_{5y}.B + \sigma_{5v}.\overline{C} + \overline{\sigma}_{5v}.2C + \sigma_{6y}.2B$$

$$+ \overline{\sigma}_{6y}.2A + \sigma_{6v}.(\overline{B} + 2C) + \overline{\sigma}_{6v}.B.$$

$$\underline{\underline{T}}\ \eta_1\zeta_1^2 : \theta_y.\overline{T}_9 + \overline{\theta}_y.(T_9 + 2\overline{T}_6 + 2\overline{T}_8) + \theta_v.6\overline{T}_{10} + \overline{\theta}_v.6\overline{T}_9 + \overline{\beta}_y(\overline{S}_5 + S_6)$$

$$+ \overline{\beta}_V.\ 5\overline{S}_6 + \gamma_y.\overline{S}_5 + \overline{\gamma}_y.\ (S_5 + 2\overline{S}_3 + 2\overline{S}_4) + \gamma_v.4S_6 + \overline{\gamma}_v.4\overline{S}_5$$

$$+ \overline{\sigma}_{5y}.(C + \overline{B}) + \overline{\sigma}_{5v}.3\overline{C} + \sigma_{6y}.\overline{B} + \overline{\sigma}_{6y}.(B + 2\overline{A}) + \sigma_{6v}.2\overline{C} + \overline{\sigma}_{6v}.2\overline{B}.$$

$$\underline{\underline{S}}\ \zeta_1^3 : \quad \theta_y.T_{10} + \overline{\theta}_y.T_9 + \theta_v.\overline{T}_{10} + \overline{\theta}_v.6T_{10} + \gamma_y.S_6 + \overline{\gamma}_y.S_5 + \gamma_v.\overline{S}_6$$

$$+ \overline{\gamma}_v.4S_6 + \sigma_{6y}.C + \overline{\sigma}_{6y}.B + \sigma_{6v}.\overline{C} + \overline{\sigma}_{6v}.2C.$$

$$\underline{\underline{T}}\ \zeta_1^3 : \quad \overline{\theta}_y.(T_{10} + \overline{T}_9) + \overline{\theta}_v.\ 7\overline{T}_{10} + \overline{\gamma}_y.(S_6 + \overline{S}_5) + \overline{\gamma}_v.5\overline{S}_6 + \overline{\sigma}_{6y}.\ (C + \overline{B})$$

$$+ \overline{\sigma}_{6v}.3\overline{C}.$$

APPENDIX 4

a- and b- coefficients variation with pupil plane position

The effect of a change δp in p, the pupil plane position, can be expressed in a form corresponding to a simple transformation of coordinates - (see equation. (36)).

$$\underline{\underline{S}} = {}^*\underline{\underline{S}} + X{}^*\underline{\underline{T}}\ ,\ \underline{\underline{T}} = \underline{\underline{T}} \tag{A1}$$

$$\text{where } X = l_{01}\cdot\left(\frac{1}{1-{}^*p/l_{01}} - \frac{1}{1-p/l_{01}}\right)$$

By substituting for $\underline{\underline{S}}$, $\underline{\underline{T}}$, ξ, η and ζ in the aberration expansion using A1 and grouping terms ${}^*\underline{\underline{S}}\ {}^*\xi^i{}^*\eta^j{}^*\zeta^k$ and ${}^*\underline{\underline{T}}{}^*\xi^i{}^*\eta^j{}^*\zeta^k$ the relations between the new and old coefficients may be found. Noting that any coefficient $\overset{(n)}{G}\mu\nu$ is linearly independent on s_o, t_o (since $\Delta\ \underline{\underline{A}} = \Delta N\ (s_o\ \underline{\underline{T}} - t_o\ \underline{\underline{S}})$, the a- and b- coefficients may be found from the expressions for the ${}^*\overset{(n)}{G}\mu\nu$.

$$\begin{aligned} \overset{(n)}{G}\mu\nu &= \overset{(n)}{G}\mu\nu a\ s_o + \overset{(n)}{G}\mu\nu b\ \ t_o \\ &= \overset{(n)}{G}\mu\nu a.({}^*s_o + X{}^*t_o) + \overset{(n)}{G}\mu\nu b.{}^*t_o \end{aligned}$$

$$\text{so}\quad \overset{(n)}{G}\mu\nu = \overset{(n)}{G}\mu\nu a.\ {}^*s_o + (\overset{(n)}{G}\mu\nu a\ X + \overset{(n)}{G}\mu\nu b){}^*\ t_o - - - - \tag{A2}$$

The a- coefficients are the coefficients of the terms $s_o\ \underline{\underline{S}}\ \xi^i\ \eta^j\ \zeta^k$, and $S_o\ \underline{\underline{T}}\ \xi^i\ \eta^j\ \zeta^k$ and the corresponding starred quantities, while the b- coefficients are the coefficients of the terms $t_o\ \underline{\underline{S}}\ \xi^i\ \eta^j\ \zeta^k$ and $t_o\ \underline{\underline{T}}\ \xi^i\ \eta^j\ {}^1\zeta^k$.

Buchdahl gives the primary a- and b- transformations and the secondary a- transformations. In the following the details of deriving the primary a- and b- transformations will be given and then the complete set of secondary and tertiary a- and b- transformations.

$$\underline{\underline{\varepsilon}} = A\,\underline{\underline{S}}\,\xi + \overline{A}\,\underline{\underline{T}}\,\xi + B\,\underline{\underline{S}}\,\eta + \overline{B}\,\underline{\underline{T}}\eta + C\,\underline{\underline{S}}\zeta + \overline{C}\,T\,\zeta + 0(5).. \quad (A3)$$

or
$$\underline{\underline{\varepsilon}} = {*A}{*\underline{\underline{S}}}{*\xi} + {*\overline{A}}\,{*\underline{\underline{T}}}\,{*\xi} + {*B}{*\underline{\underline{S}}}{*\eta} + {*\overline{B}}{*\underline{\underline{T}}}{*\eta} + {*C}{*\underline{\underline{S}}}{*\zeta} + {*\overline{C}}{*\underline{\underline{T}}}{*\zeta} + 0(5) \quad (A4)$$

Substituting in A3 using A1 gives

$$\underline{\underline{\varepsilon}} = A({*\underline{\underline{S}}} + X{*T})({*\xi} + 2X{*\eta} + X^2{*\zeta}) + \overline{A}{*T}\,({*\xi} + 2X{*\eta} + X^2{*\zeta})$$

$$+ B({*S} + X{*\underline{\underline{T}}})\,({*\eta} + X{*\zeta}) + \overline{B}\,\underline{\underline{T}}\,({*\eta} + X{*\zeta}) + C({*\underline{\underline{S}}} + X\underline{\underline{T}}){*\zeta} + C{*\underline{\underline{T}}}{*\zeta}$$

$$= A{*\underline{\underline{S}}}{*\xi} + (\overline{A} + AX){*\underline{\underline{T}}}{*\xi} + (2AX + B){*\underline{\underline{S}}}{*\eta} + (2AX^2 + (2\overline{A} + B)\,X$$

$$+ B){*\underline{\underline{T}}}{*\eta} + (AX^2 + BX + C)\,{*\underline{\underline{S}}}{*\zeta} + (AX^3 + (\overline{A} + B)\,X^2 + (\overline{B} + C)X$$

$$+ \overline{C})\,*\,\underline{\underline{T}}{*\zeta}$$

Hence ${*A} = A$, ${*\overline{A}} = \overline{A} + AX$

$$(A5)$$

$${*B} = 2AX + B, \qquad {*\overline{B}} = 2AX^2 + (2\overline{A} + B)X + B$$

$${*C} = AX^2 + BX + C, \qquad {*\overline{C}} = AX^3 + (\overline{A} + B)\,X^2 + (\overline{B} + C)\,X + \overline{C}.$$

By consideration of A2 it is possible to read off from A5 the a- and b- coefficients, so

$$\delta A_a = 0,\ \delta\overline{A}_a = A_a X,\ \delta B_a = 2A_a X,\ \delta\overline{B}_a = 2A_a X^2 + (2\overline{A}_a + B_a)\,X$$

$$\delta C_a = A_a X^2 + B_a X,\ \delta\overline{C}_a = A_a X^3 + (\overline{A}_a + B_a)\,X^2 + (\overline{B}_a + C_a)\,X$$

Now $X = l_{01}\left(\frac{1}{1-{}^*p/l_{01}} - \frac{1}{1-p/l_{01}}\right) = \delta p/[(1-{}^*p/l_{01})(1-p/l_{01})]$

so the derivatives of the coefficients with respect to pupil change depend only on terms of first-order in X in the transformation expressions S_o

$$\frac{\partial A_a}{\partial p} = 0, \frac{\partial \bar{A}_a}{\partial p} = A_a\psi , \frac{\partial B_a}{\partial p} = 2A_a\psi , \frac{\partial \bar{B}_a}{\partial p} = (2\bar{A}_a + B_a)\psi ,$$

$$\frac{\partial C_a}{\partial p} = B_a\psi , \frac{\partial \bar{C}_a}{\partial p} = (\bar{B}_a + C_a)\Psi , \text{ where } \psi^{-1} = (1 - p/l_{01})^2 \tag{A6}$$

For the b- coefficients,

$${}^*A_b = A_aX + A_b , \quad {}^*\bar{A}_b = (\bar{A}_a + A_aX)\, X + \bar{A}_b + A_bX$$

$${}^*B_b = (2A_aX + B_a)\, X + (2A_b\, X + B_b), \; {}^*B_b = 2A_aX^3 + (2A_b + 2\bar{A}_a + B_a)X^2$$

$$+ (2\bar{A}_b + B_b + B_a)\, X + B_b$$

and so on, so that

$$\frac{\partial A_b}{\partial p} = A_a\psi, \frac{\partial \bar{A}_b}{\partial p} + (\bar{A}_a + A_b)\psi , \frac{\partial B_b}{\partial p} = (B_a + 2A_b)\psi$$

$$\frac{\partial \bar{B}_b}{\partial p} = (2\bar{A}_b + B_b + \bar{B}_a)\psi , \frac{\partial C_b}{\partial p} = (C_a + B_b)\psi , \frac{\partial \bar{C}_b}{\partial p} = (\bar{C}_a + \bar{B}_b + C_b)\psi \tag{A7}$$

The secondary and tertiary derivatives are obtained by the above method and are given below.

$$\frac{\partial S_{1a}}{\partial p} = 0 \qquad\qquad \frac{\partial \bar{S}_{1a}}{\partial p} = S_{1a}\psi$$

$$\frac{\partial S_{2a}}{\partial p} = 4S_{1a}\psi \qquad \frac{\partial \bar{S}_{2a}}{\partial p} = (4\bar{S}_{1a} + S_{2a})\psi$$

$$\frac{\partial S_{2a}}{\partial p} = S_{2a}\psi \qquad \frac{\partial \bar{S}_{3a}}{\partial p} = (\bar{S}_{2a} + S_{3a})\psi$$

$$\frac{\partial S_{4a}}{\partial p} = 2S_{2a}\psi \qquad \frac{\partial \bar{S}_{4a}}{\partial p} = (2\bar{S}_{2a} + S_{4a})\psi$$

$$\frac{\partial S_{5a}}{\partial p} = 2(S_{3a} + S_{4a})\psi \qquad \frac{\partial \bar{S}_{5a}}{\partial p} = (2\bar{S}_{3a} + 2\bar{S}_{4a} + S_{5a})\psi$$

$$\frac{\partial S_{4a}}{\partial p} = S_{5a}\psi \qquad \frac{\partial \bar{S}_{6a}}{\partial p} = (\bar{S}_{5a} + S_{6a})\psi$$

$$\frac{\partial S_{1b}}{\partial p} = S_{1a}\psi \qquad \frac{\partial \bar{S}_{1b}}{\partial p} = (\bar{S}_{1a} + S_{1b})\psi$$

$$\frac{\partial S_{2b}}{\partial p} = (S_{2a} + 4S_{1b})\psi \qquad \frac{\partial \bar{S}_{2b}}{\partial p} = (\bar{S}_{2a} + 4\bar{S}_{1b} + S_{2b})\psi$$

$$\frac{\partial S_{3b}}{\partial p} = (S_{3a} + S_{2b})\psi \qquad \frac{\partial \bar{S}_{3b}}{\partial p} = (\bar{S}_{3a} + \bar{S}_{2b} + S_{3b})\psi$$

...(A9)

$$\frac{\partial S_{4b}}{\partial p} = (S_{4a} + 2S_{2b})\psi \qquad \frac{\partial \bar{S}_{4b}}{\partial p} = (\bar{S}_{4a} + 2\bar{S}_{2b} + S_{4b})\psi$$

$$\frac{\partial S_{5b}}{\partial p} = (S_{5a} + 2S_{3b} + 2S_{4b})\psi \qquad \frac{\partial \bar{S}_{5b}}{\partial p} = (\bar{S}_{5a} + 2\bar{S}_{3b} + 2\bar{S}_{4b} + S_{5b})\psi$$

$$\frac{\partial S_{6b}}{\partial p} = (S_{6a} + S_{5b})\psi \qquad \frac{\partial \bar{S}_{6b}}{\partial p} = (\bar{S}_{6a} + \bar{S}_{5b} + S_{6b})\psi$$

and for the tertiary coefficients.

$$\frac{\partial T_{1a}}{\partial p} = 0 \qquad \frac{\partial \bar{T}_{1a}}{\partial p} = T_{1a}\cdot\psi$$

$$\frac{\partial T_{2a}}{\partial p} = 6T_{1a}\cdot\psi \qquad \frac{\partial \bar{T}_{2a}}{\partial p} = (6\bar{T}_{1a} + T_{2a})\psi$$

$$\frac{\partial T_{3a}}{\partial p} = T_{2a}\cdot\psi \qquad \frac{\partial \bar{T}_{3a}}{\partial p} = (\bar{T}_{2a} + T_{3a})\psi$$

$$\frac{\partial T_{4a}}{\partial p} = 4T_{2a}\psi \qquad \frac{\partial \bar{T}_{4a}}{\partial p} = (4\bar{T}_{2a} + T_{4a})\psi$$

$$\frac{\partial T_{5a}}{\partial p} = 2(2T_{3a} + T_{4a})\psi \qquad \frac{\partial \bar{T}_{5a}}{\partial p} = (4\bar{T}_{3a} + 2\bar{T}_{4a} + T_{5a})\psi \qquad \text{-----} \quad \text{(A10)}$$

$$\frac{\partial T_{6a}}{\partial p} = T_{5a}\psi \qquad \frac{\partial \bar{T}_{6a}}{\partial p} = (\bar{T}_{5a} + T_{6a})\psi$$

$$\frac{\partial T_{7a}}{\partial p} = 2T_{4a}\psi \qquad \frac{\partial \bar{T}_{7a}}{\partial p} = (2\bar{T}_{4a} + T_{7a})\psi$$

$$\frac{\partial T_{8a}}{\partial p} = (2T_{5a} + 3T_{7a})\psi \qquad \frac{\partial \bar{T}_{8a}}{\partial p} = (2\bar{T}_{5a} + 3\bar{T}_{7a} + T_{8a})\psi$$

$$\frac{\partial T_{9a}}{\partial p} = 2(T_{6a} + T_{8a})\psi \qquad \frac{\partial \bar{T}_{9a}}{\partial p} = (2\bar{T}_{6a} + 2\bar{T}_{8a} + T_{9a})\psi$$

$$\frac{\partial T_{10a}}{\partial p} = T_{9a}\psi \qquad \frac{\partial T_{10a}}{\partial p} = (\bar{T}_{9a} + T_{10a})\psi$$

The derivatives of the tertiary ba- coefficients are not required.

Appendix 5

The following pages give listings of the programs in which the various assessment functions are calculated. The listings have been given to facilitate the programming of these functions for others who may wish to carry out similar computations, and also to provide complete expressions for some of the functions described in the body of this thesis. It will be noted that the programs contain a small amount of machine code and file-handling procedures; these are retained in the listing to indicate that the programs are part of an assessment suite rather than being complete in themselves. Also some idea of the file-handling involved may be determined.

```
P5B,    30-8-74;
begin integer i,n,j,volume;          comment CBS:;     array oddbits[1:200];     integer array ODDS,OPTXD[1:15];
      real r1,pp,K;
          procedure chain(program);  value program;  integer program;
          comment calls & enters a program dumped on disk by AFDUMP. Program
                is the program name in rappacked form;
               code 30 9 : 46 1,
                    40 0, : 26 7886
                    72 256 : 30 <+21>
                    72 6659 : 30 <-47>
                    20 11, : 75 6656
                    32 11, : 45 4,
                    26 7913 : 30 program
                    50 24 : 72 6657
                    54 12 : 72 6657
                    54 12 : 72 6657
                    44 8166 : 00 0
                    00 0 : 00 0;
          comment end of  chain;

          procedure change;
               begin oddbits[98]:=(if j=3 and oddbits[200]<0.1 then 0 else if j=1 then 1 else 2);
                    oddbits[99]:=(if j=3 and oddbits[200]<0.1 then 0 else  -1);
                    oddbits[100]:=(if j=3 and oddbits[200]<0.1 then 0 else if j=1 then 0 else -1);
                    oddbits[90]:=(if j=2 then 0.25 else 0.5)*oddbits[90];
               end change;

          volume:=(if ng(18) then 19 else 8);
          clear(ODDS);    code 30 <0ODDS>: 20 i;         ODDS[1]:=i;    ODDS[2]:=3;    ODDS[7]:=1;
          ODDS[8]:=volume;          open(ODDS);          ODDS[10]:=size(oddbits);          O
          ODDS[13]:=address(oddbits);          in40(ODDS);
          n:=entier(oddbits[1]);  r1:=oddbits[2];          pp:=oddbits[3];

begin     real a2,a4,a6,a0,b1,b3,b5,b7,c2,c4,c6,d3,d5,e4,as0,as2,as4,as6,bs1,bs3,bs5,bs7,cs2,cs4,cs6,ds3,ds5,ds7,
          es4,es6,fs5,fs7,gs6,hs7,V,V2 V4,a,b,c,br1,br3,br5,br7,cr2 cr4,cr6,dr3,dr5,er4,bsq,csq,chi,asq,bsq2,
c         csq2,abk2,abk4,abk6,bk5,chibr,crk4,brs1,brs3 brs5,brs7,crs2 crs4,crs6,drs3,drs5,drs7,ers4,ers6,frs5,
          frs7,grs6,hrs7,brk5,alf0,alf2,alf4,alf6,alf8,Assym,Eysq,Ezsq,Y0,Y2 Y4,Y6,Y8,Y10,Y12,Y14,Z0,Z2,Z4,Z6,
          Z8,Z10,Z12,Z14,RG,Ast,bk3,ck4,X1,X3,X5 X7,V1,V3,V5,V7,I4 I5,xdash,ya,va,ia,yb,vb,ib,cc,kk,q,omega,ap,
          apb,bp,bpb,cp,cpb,vpd,aq,aqb,bq,bqb,cq,cqb,N1,qq,da,db,dc,sum1,sum2,sum3,P11,P12,P13,P14,P15,P16,P21,
          P22,P23,P24,P25,P26,weight,xd,intxd;
          integer i,angno,ii,jj,p,m,sb,sur,apno,shifno,app,ang,shif;
```

```
array sig[1:5],mu[1:12],tau[1:20],rhosb,radj,Y,Z[0:10];
comment CBS:;  array cur,d,k,N,rho[1:n+1],yp,vp[1:2,1:n+1],P1,P2,P3[1:5,1:5,1:20],xopt[1:2,1:5,1:5];
comment CBS:;   array coeff[1:156*n+2],system[1:7*n+1],deriv[1:4*n*(n+15)],dPtotdc[1:2,1:6,1:n],
     dStotdc[1:2,1:12,1:n],dTtotdc[1:2,1:20,1:n],A,AB[0:1,1:3],S,SB[0:1,1:6],T[0:1,1:10],temp[1:200],
     costab,sintab[-1:10];
integer array ABCO,DERIV,SYS MISC[1:15],vig[1:3];
switch L:=more,finish,newplane,crossover,minfound,checkmin;

procedure trace;
     begin yp[1,1]:=1.0;   vp[1,1]:=r1;    vp[2,1]:=1.0/(1.0-pp*r1);   yp[2,1]:=pp*vp[2,1];
          for p:=1,2 do for i:=1 step 1 until n do
               begin yp[p,i]:=if i≠1 then yp[p,i-1]-d[i]*vp[p,i] else yp[p,i];
                    vp[p,i+1]:=vp[p,i]+(cur[i]*yp[p,i]-vp[p,i])*(1.0-k[i]);
               end i,p;
     end trace;

real procedure bisec(f,x,a,b,eps,error);  value a,b,eps;  real f,x,a,b,eps;  label error;
     begin real q;  x:=a;  q:=f;  x:=b;
          if f*q>0 then goto error;
          q:=(b-a)*sign(q)/2.0;  x:=(a+b)/2.0;  eps:=eps/2.0;
          for q:=q/2.0 while abs(q)>eps do x:=x+(if f>0 then q else -q);

          bisec:=x
     end bisec;

procedure pupilit(a,vig,rhosb,V,tol);  value a,V,tol;  real a,V,tol;  arrayrhosb;  integer array vig;
     begin comment segment:;
     real theta,Sy,Sz,Ty,xi,eta,zeta,xi2,eta2,chi,eps,rad,rad2,ya,yb,va,vb,ya2,yava,yavb,yayb,ybva,
          yb2,ybvb,brack,rho0,csb,coz,zin;
     integer i,j,jj,ii,num;              array rhosur[0:10];
     switch L:=iterate;
     for j:=0 step 1 until 10 do rhosb[j]:=rho[1];   rho0:=rho[1];
     for i:=1 step 1 until a do
          begin j:=vig[i];
               ya:=yp[1,j];  yb:=yp[2,j];
               va:=vp[1,j];  vb:=vp[2,j];
               yava:=ya*va;  yavb:=ya*vb;
               ybva:=yb*va;  ybvb:=yb*vb;
               ya2:=ya*ya;   yb2:=yb*yb;    yayb:=ya*yb;
               brack:=yavb+ybva;  csb:=cur[j];  ii:=-2;
               zeta:=V*V;  rad:=rho[j];  rad2:=rad*rad;
```

```
                    for theta:=0.0 step 0.3142 until 3.0,3.1415927 do
                        begin ii:=ii+1;   jj:=0;        coz:=costab[ii];         zin:=sintab[ii];
         iterate:               Sy:=rho0*coz;  Sz:=rho0*zin;        jj:=jj+1;
                                xi:=Sy*Sy+Sz*Sz;  eta:=Sy*V;
                                xi2:=ya2*xi+2.0*yayb*eta+yb2*zeta;
                                eta2:=yava*xi+brack*eta+ybvb*zeta;
                                chi:=xi2/(1.0+csb*eta2);
                                eps:=sqrt(chi)-rad;
                                if eps>tol*rad or eps<-tol*rad then
                                    begin rho0:=rho0-0.75*eps/ya;  goto iterate end;
                                if ii=-1 then
                                    begin rho0:=1.1*rho0;          ii:=jj:=0;  goto iterate    end;
                                rhosur[ii]:=rho0;  rho0:=1.1*rho0;
                        end theta;
                    for j:=0 step 1 until 10 do
                        begin if rhosur[j]<rhosb[j] then rhosb[j]:=rhosur[j] end;
                end i;
      end procedure;

procedure fitellipse(y,z,a,b,c);  array y,z;  real a,b,c;
    begin comment segment:; real p1,p2,p3,p4,dela,delb,delc,Y,Z a1,a2,a3,b1,b2,b3,c1,c2,c3,d1,d2,d3,den,a0,
          b0,c0,q1,q2,r1,r2,t1,t2;
          integer j,count;  switch fit:=again;
          a0:=(y[0]-y[10])*0.5;  b0:=z[5];  c0:=(y[0]+y[10])*0.5;    count:=0;
again:    a1:=a2:=a3:=b1:=b2:=b3:=c1:=c2:=c3:=d1:=d2:=d3:=0.0;     count:=count+1;
          for j:=0 step 1 until 10 do
              begin Y:=y[j]-c0;  Z:=z[j];
                    p1:=2.0*Y*Y/(a0*a0*a0);
                    p2:=2.0*Z*Z/(b0*b0*b0);
                    p3:=2.0*Y/(a0*a0);
                    p4:=1.0-p1*a0*0.5-p2*b0*0.5;
                    a1:=a1+p1*p1;  a2:=a2+p1*p2;  a3:=a3+p1*p3;
                    b1:=b1+p1*p2;  b2:=b2+p2*p2;  b3:=b3+p2*p3;
                    c1:=c1+p3*p1;  c2:=c2+p2*p3;  c3:=c3+p3*p3;
                    d1:=d1+p4*p1;  d2:=d2+p2*p4;  d3:=d3+p3*p4;
              end j;
          q1:=a1*c2-a2*c1;  q2:=a1*c3-a3*c1;
          r1:=b1*c2-b2*c1;  r2:=b1*c3-b3*c1;
          t1:=d1*c2-d2*c1;  t2:=d1*c3-d3*c1;
          den:=q1*r2-q2*r1;
          if abs(den)<₁₀-5 then print S$14t2?ILL-CONDITIONED EQUATIONS$122;
```

```
        dela:=(t1*r2-t2*r1)/den;
        delb:=(t1-dela*q1)/r1;
        if abs(c1)<₁₀-7 then print ££l??C1 =?,c1,t1-dela*q1,c2,c3;
        delc:=(d1-a1*dela-b1*delb)/c1;
        if count<3 then begin a0:=a0-dela;   b0:=b0-delb;   c0:=c0-delc;    goto again  end;
        a:=a0-dela;  b:=b0-delb;  c:=c0-delc;
    end procedure fitellipse;

        procedure Effocab(sig,mu,tau);          array sig,mu,tau;
          comment Has global variables mu[1:12],sigma[1:5] as well as usual A,AB S,SB arrays and K;
            begin comment segment; real KK;
                  KK:=0.5*K;
                  sig[1]:=A[0,1]*K;    sig[2]:=AB[0,1]*K;    sig[3]:=AB[0,2]*KK;
                  sig[4]:=A[0,3]*K-sig[3];     sig[5]:=AB[0,3]*K;
                  mu[1]:=S[0,1]*K;  mu[3]:=S[0,2]*KK;   mu[2]:=K*SB[0,1]+mu[3];   mu[5]:=S[0,3]*K;
                  mu[4]:=SB[0,2]*K+mu[5];   mu[6]:=S[0,4]*K;   mu[8]:=(SB[0,4]+S[0,5])*KK;
                  mu[7]:=SB[0,3]*K+mu[8];   mu[9]:=S[0,5]*KK;   mu[10]:=(SB[0,5]+S[0,6])*K;
                  mu[11]:=S[0,6]*K;   mu[12]:=SB[0,6]*K;

                  tau[1]:=T[0,1]*K;    tau[3]:=T[0,2]*KK;    tau[2]:=K*T[1,1]+tau[3];
                  tau[5]:=T[0,3]*K;    tau[4]:=T[1,2]*K+tau[5];    tau[6]:=T[0,4]*K;
                  tau[7]:=(T[1,3]+0.5*T[1,4]+0.5*T[0,5]+0.375*T[0,7])*K;
                  tau[9]:=(0.5*T[0,5]+0.25*T[0,7])*K;    tau[8]:=KK*(T[1,4]+T[0,5]+T[0,7]);
                  tau[10]:=0.125*T[0,7]*K;    tau[11]:=(T[1,5]+T[0,6])*K;    tau[12]:=(T[1,7]+T[0,8])*K;
                  tau[13]:=T[0,6]*K;    tau[14]:=T[0,8]*K;    tau[15]:=(T[1,6]+0.5*(T[1,8]+T[0,9]))*K;
                  tau[16]:=KK*(T[1,8]+T[0,9]);    tau[17]:=KK*T[0,9];    tau[18]:=(T[1,9]+T[0,10])*K;
                  tau[19]:=T[0,10]*K;    tau[20]:=T[1,10]*K;
            end Effocab;

        real procedure best focus(ro,H);      value ro,H;            real ro,H;
            begin integer i,j;                real xstar,ro2,H2;     array x[1:3,0:3];
                  x[1,0]:=0.33333333*sig[1];
                  x[1,1]:=sig[3]+0.5*sig[4];
                  x[2,0]:=0.25*mu[1];
                  x[2,1]:=0.1666667*(mu[4]+mu[5]+mu[6]);
                  x[2,2]:=0.25*(mu[10]+mu[11]);
                  x[3,0]:=0.20*tau[1];
                  x[3,1]:=0.125*(tau[4]+tau[5]+tau[6]);
                  x[3,2]:=0.166666667*(tau[11]+tau[13]+0.75*tau[12]+0.25*tau[14]);
                  x[3,3]:=0.250*(tau[18]+tau[19]);
```

```
            xstar:=0.0;             ro2:=ro*ro;             H2:=H*H;
            for i:=1 step 1 until 3 do for j:=0 step 1 until i do
                xstar:=xstar+x[i,j]*ro2↑(i-j)*(if abs(H2)<₁₀-8 and j=0 then 1.0 else H2↑j);
            bestfocus:=-2.0*xstar;
        end best focus;

    prefix($$s4??);  scaled(5);  digits(2);         lineprinter;

    sb:=entier(oddbits[4]);    angno:=entier(oddbits[5]);

    begin comment segment:;       real dum2;
    clear(ABCO);  clear(DERIV);  clear(SYS);
    code 30<9SPECS>:20i;  SYS[1]:=i;
    code 30<9ABS>:20i;  ABCO[1]:=i;
    SYS[2]:=ABCO[2]:=DERIV[2]:=3;
    SYS[7]:=ABCO[7]:=DERIV[7]:=1;
    SYS[8]:=ABCO[8]:=DERIV[8]:=volume;
    open(SYS);  open(ABCO);
    ABCO[10]:=size(coeff);  ABCO[13]:=address(coeff);  in40(ABCO);
    SYS[10]:=size(system);  SYS[13]:=address(system);  in40(SYS);
    for m:=0,1 do
        begin for j:=1,2,3 do
                begin ii:=156*(n-1)+78*m;
                      A[m,j]:=coeff[ii+4*j];  AB[m,j]:=coeff[ii+4*j+2]
                end j;
            for j:=1 step 1 until 6 do
                begin ii:=156*(n-1)+78*m+12;
                      S[m,j]:=coeff[ii+4*j];  SB[m,j]:=coeff[ii+4*j+2]
                end j;
            for j:=1 step 1 until 10 do
                begin ii:=156*(n-1)+36+2*m;  T[m,j]:=coeff[ii+4*j] end j;
            for j:=1 step 1 until n do
                begin yp[1+m,j]:=coeff[156*(j-1)+78*m+1];
                      vp[1+m,j]:=coeff[156*(j-1)+78*m+2];
                 end j;
             vp[1+m,n+1]:=coeff[156*n+m+1];
        end m;
    for j:=1 step 1 until n do
        begin cur[j]:=system[7*(j-1)+1];     d[j]:=system[7*j-2];
              N[j]:=system[7*j-1];    rho[j]:=system[7*j];
        end;
```

```
N[n+1]:=system[7*n+1];  for j:=1 step 1 until n do k[j]:=N[j]/N[j+1];
end dum2;
apno:=entier(oddbits[40]+0.01);    shifno:=entier(oddbits[46]+0.01);
K:=1.0;      Effocab(sig,mu,tau);
jj:=0;    clear(temp);  sb:=entier(oddbits[31]+0.01);
sum1:=sum2:=sum3:=0.0;
j:=-2;
              for app:=1 step 1 until apno do
                  begin a:=b:=oddbits[40+app];                c:=0.0;      jj:=0;;
more:   jj:=jj+1;    ang:=jj;    V:=oddbits[5+jj];  if V>90 then goto finish;
V:=tan(V*0.01745329);  V2:=V*V;  V4:=V2*V2;
a0:=if ng(25) then 0.0 else (sig[5]+mu[12]*V2+tau[20]*V4)*V2*V;
a2:=(sig[2]+(mu[7]-mu[8])*V2+(tau[15]-tau[16])*V4)*V;
a4:=((mu[2]-mu[3])+(tau[7]-tau[8]+tau[10])*V2)*V;
a6:=(tau[2]-tau[3])*V;
b1:=(3.0*sig[3]+sig[4]+mu[10]*V2+tau[18]*V4)*V2;
b3:=(sig[1]+mu[4]*V2+tau[11]*V4);
b5:=mu[1]+tau[4]*V2;
b7:=tau[1];
c2:=(sig[2]+mu[8]*V2+tau[16]*V4)*2.0*V;
c4:=2.0*(mu[3]+(tau[8]-4.0*tau[10])*V2)*V;
c6:=2.0*tau[3]*V;
d3:=(mu[6]+tau[12]*V2)*V2;
d5:=tau[6]*V2;  e4:=8.0*tau[10]*V2*V;
br1:=(sig[3]+sig[4]+mu[11]*V2+tau[19]*V4)*V2;
br3:=sig[1]+mu[5]*V2+tau[13]*V4;
br5:=mu[1]+tau[5]*V2;  br7:=tau[1];
cr2:=2.0*(sig[2]+mu[9]*V2+tau[17]*V4)*V;
cr4:=2.0*(mu[3]+(tau[9]-2.0*tau[10])*V2)*V;
cr6:=2.0*tau[3]*V;
dr3:=(mu[6]+tau[14]*V2)*V2;  dr5:=tau[6]*V2;
er4:=e4;
csq:=c*c;  bsq:=b*b;  chi:=a*a-bsq;
abk2:=a2+c*b3+csq*c4+csq*c*d5;
abk4:=a4+c*b5+csq*c6;
abk6:=a6+c*b7;
bk3:=b3+2.0*c*c4+3.0*csq*d5;
bk5:=b5+2.0*c*c6;
chibr:=2.0*a*a+chi;
```

```
asq:=a*a;  bsq2:=bsq*bsq;  csq2:=csq*csq;
as0:=a0+c*b1+csq*c2+csq*(c*d3+csq*e4)+abk2*csq+abk4*csq2+abk6*csq2*csq;
as2:=abk2*bsq+abk4*2.0*bsq*csq+abk6*3.0*bsq*csq2;
as4:=abk4*bsq2+abk6*3.0*bsq2*csq;
as6:=abk6*bsq2*bsq;
bs1:=a*(b1+2.0*c*c2+3.0*csq*d3+4.0*csq*c*e4+abk2*2.0*c+csq*bk3+4.0*csq*c*abk4+csq2*bk5
      +6.0*csq2*c*abk6+b7*csq2*csq);
bs3:=a*(bk3*bsq+4.0*bsq*c*abk4+bk5*2.0*bsq*csq+abk6*12.0*bsq*csq*c+b7*3.0*bsq*csq2);
bs5:=a*(bk5*bsq2+abk6*6.0*bsq2*c+b7*3.0*bsq2*csq);
bs7:=a*b7*bsq2*bsq;  ck4:=c4+3.0*c*d5;
cs2:=abk2*chi+bk3*2.0*asq*c+asq*csq*ck4+abk4*csq*2.0*chibr+4.0*asq*csq*c*bk5+
          asq*csq2*c6+abk6*3.0*csq2*(4.0*asq+chi)+asq*(6.0*csq2*c*b7+c2+3.0*c*d3+6.0*csq*e4);
cs4:=bsq*(2.0*chi*abk4+4.0*asq*c*bk5+2.0*asq*csq*c6+6.0*csq*(chibr*abk6+2.0*asq*c*b7));
cs6:=bsq2*3.0*(chi*abk6+2.0*asq*c*b7);
ds3:=a*(4.0*chi*c*abk4+2.0*chibr*csq*bk5+a*c6*4.0*a*csq*c+abk6*4.0*csq*c*(3.0*chi+2.0*asq)
          +b7*3.0*(2.0*chibr-chi)*csq2+chi*bk3+2.0*asq*c*ck4+asq*csq*d5+4.0*asq*c*e4+asq*d3);
ds5:=a*bsq*(2.0*chi*bk5+4.0*asq*c*c6+12.0*chi*c*abk6+6.0*chibr*csq*b7+asq*d5);
ds7:=3.0*a*bsq2*chi*b7;
es4:=3.0*chi*csq*(4.0*asq+chi)*abk6+4.0*asq*csq*c*b7*(3.0*chi+2.0*asq)+chi*chi*abk4
          +asq*(4.0*chi*c*bk5+2.0*chibr*csq*c6+chi*ck4+2.0*asq*c*d5+asq*e4);
es6:=3.0*bsq*chi*(abk6*chi+4.0*asq*c*b7)+2.0*asq*bsq*chi*c6;
fs5:=a*chi*(6.0*chi*c*abk6+chi*b5+3.0*b7*csq*(4.0*asq+chi)+4.0*asq*c*c6+2.0*chi*c*c6+asq*d5);
fs7:=3.0*a*bsq*chi*chi*b7;
gs6:=chi*chi*(a6*chi+b7*6.0*asq*c+c*chi*b7+asq*c6);
hs7:=a*chi*chi*chi*b7;

crk4:=cr4+cr6*csq;  brk5:=br5+c*cr6;
brs1:=b*(br1+csq*(br3+csq*(br5+csq*br7))+c*(cr2+csq*cr4
      +csq2*cr6)+csq*(dr3+csq*dr5)+csq*c*e4);
brs3:=bsq*b*(br3+2.0*csq*br5+3.0*csq2*br7+c*(cr4
      +2.0*csq*cr6)+csq*dr5);
brs5:=bsq2*b*(br5+3.0*csq*br7+c*cr6);
brs7:=bsq2*bsq*b*br7;
crs2:=a*b*(2.0*c*br3+4.0*csq*c*br5+6.0*c*csq2*br7+(cr2
      +csq*crk4)+2.0*c*c*crk4+c*(4.0*csq*dr5+2.0*dr3+3.0*c*er4));
crs4:=bsq*b*a*(4.0*c*(br5+3.0*csq*br7)+crk4+2.0*c*dr5);
crs6:=bsq2*b*a*(c*6.0*br7+cr6);
drs3:=b*(chi*(br3+c*cr4+csq*dr5)+2.0*asq*c*(cr4+2.0*c*dr5)
      +asq*csq*dr5+brk5*csq*2.0*chibr+4.0*asq*csq*c*cr6
      +3.0*csq2*(4.0*asq+chi)*br7+asq*(dr3+3.0*c*e4));
```

```
drs5:=bsq*b*(brk5*2.0*chi+4.0*asq*c*cr6+6.0*csq*chibr*br7+asq*dr5);
drs7:=3.0*bsq2*b*chi*br7;
ers4:=a*b*(4.0*brk5*c*chi+4.0*csq*c*br7*(3.0*chi+2.0*asq)
        +2.0*chibr*csq*cr6+(cr4+2.0*c*dr5)*chi+asq*(2.0*c*dr5+e4));
ers6:=2.0*a*bsq*b*chi*(c*6.0*br7+cr6);
frs5:=b*chi*(brk5*chi+4.0*asq*c*cr6+3.0*csq*(4.0*asq+chi)*br7+asq*dr5);
frs7:=3.0*bsq*b*chi*chi*br7;  hrs7:=b*chi*chi*chi*br7;
grs6:=a*chi*chi*b*(6.0*c*br7+cr6);
alf0:=2.0*as0;  alf2:=2.0*as2+cs2;
alf4:=2.0*as4+cs4+0.75*es4;
alf6:=2.0*as6+cs6+0.75*es6+0.625*gs6;
Assym:=0.5*alf0+0.25*alf2+alf4/6+0.125*alf6;
Y0:=2*as0*as0;
Y2:=2.0*as0*(2.0*as2+cs2)+bs1*bs1;
Y4:=2.0*(as2*as2+2.0*as0*as4+bs1*bs3+as0*cs4+as2*cs2
        +0.375*(cs2*cs2+2.0*(as0*es4+bs1*ds3)));
Y6:=4.0*(as0*as6+as2*as4)+bs3*bs3+2.0*(bs1*bs5+as0*cs6+
        as2*cs4+as4*cs2)+1.5*(cs2*cs4+as0*es6+as2*es4+bs1*ds5
        +bs3*ds3)+0.625*(ds3*ds3+2.0*(as0*gs6+bs1*fs5+cs2*es4));
Y8:=2.0*as4*as4+4.0*as2*as6+2.0*(bs1*bs7+bs3*bs5)+2.0*(
        as2*cs6+as4*cs4+as6*cs2)+0.75*(cs4*cs4+2.0*(as2*es6+as4
        *es4+bs1*ds7+bs3*ds5+bs5*ds3))+1.25*(ds3*ds5
        +as2*gs6+bs1*fs7+bs3*fs5+cs2*es6+cs4*es4)+35/64*
        (es4*es4+2.0*(bs1*hs7+cs2*gs6+ds3*fs5));
Y10:=2.0*(2.0*as4*as6+bs3*bs7+0.5*bs5*bs5+as4*cs6+as6*cs4
        +0.75*(cs4*cs6+as4*es6+as6*es4+bs3*ds7+bs5*ds5
        +bs7*ds3)+0.3125*(ds5*ds5+2.0*(as4*gs6+bs3*fs7+
        bs5*fs5+cs4*es6+cs6*es4))+35/64*(bs3*hs7+cs4*gs6+es4*es6
        +ds3*fs7+ds5*fs5)+63/256*(fs5*fs5+2.0*(ds3*hs7+
        es4*gs6)));
Y12:=2.0*as6*as6+2.0*(bs5*bs7+as6*cs6)+0.75*(cs6*cs6+2.0*(as6
        *es6+bs5*ds7+bs7*ds5))+1.25*(ds5*ds7+as6*gs6+bs5
        *fs7+bs7*fs5+cs6*es6)+35/64*(es6*es6+2.0*(bs5*hs7+
        cs6*gs6+ds5*fs7+ds7*fs5))+63/64*(fs5*fs7+ds5*hs7+
        es6*gs6)+231/512*(gs6*gs6+2.0*fs5*hs7);
Y14:=bs7*bs7+1.5*bs7*ds7+1.25*bs7*fs7+35/32*(bs7*hs7
        +ds7*fs7)+63/128*(fs7*fs7+2.0*ds7*hs7)+0.625*ds7*ds7
        +231/512*2.0*fs7*hs7+429/1024*hs7*hs7;
Eysq:=0.5*Y0+0.25*Y2+0.166667*Y4+0.125*Y6+0.10*Y8
        +0.083333*Y10+1/14*Y12+0.0625*Y14-(if ng(24) then Assym*Assym else 0.0);
```

```
Z0:=0.0;     Z2:=brs1*brs1;
Z4:=2.0*brs1*brs3+0.25*crs2*crs2+0.5*brs1*drs3;
Z6:=brs3*brs3+2.0*brs1*brs5+0.5*crs2*crs4+0.125*drs3*drs3+0.5*(brs1*
      drs5+brs3*drs3)+0.25*(brs1*frs5+crs2*ers4);
Z8:=2.0*(brs1*brs7+brs3*brs5)+0.25*(crs4*crs4+2.0*crs2*crs6)
      +0.25*drs3*drs5+5/64*ers4*ers4+0.5*(brs1*drs7+brs3
      *drs5+brs5*drs3)+0.25*(brs1*frs7+brs3*frs5+crs2*ers6
      +crs4*ers4)+5/32*(brs1*hrs7+crs2*grs6+drs3*frs5);
Z10:=2.0*brs3*brs7+brs5*brs5+0.5*crs4*crs6+0.125*(drs5*drs5+2.0*
      drs3*drs7)+5/32*ers4*ers6+7/128*frs5*frs5+
      0.5*(brs3*drs7+brs5*drs5+brs7*drs3)+0.25*(brs3*frs7
      +brs5*frs5+crs4*ers6+crs6*ers4)+5/32*(brs3*hrs7+
      crs4*grs6+drs3*frs7+drs5*frs5)+7/64*(drs3*hrs7+ers4
      *grs6);
Z12:=2.0*brs5*brs7+0.25*(crs6*crs6+drs5*drs7)+5/64*ers6
      *ers6+7/64*frs5*frs7+21/512*grs6*grs6+0.5*(brs5*drs7
      +brs7*drs5)+0.25*(brs5*frs7+brs7*frs5+crs6*ers6)+5/32*
      (brs5*hrs7+crs6*grs6+drs5*frs7+drs7*frs5)+7/64*(drs5
      *hrs7+ers6*grs6)+21/256*frs5*hrs7;
Z14:=brs7*brs7+0.125*drs7*drs7+7/128*frs7*frs7+33/1024*
      hrs7*hrs7+0.5*brs7*(drs7+0.5*(frs7+0.625*
      hrs7))+5/32*drs7*frs7+1/64*(7.0*drs7+5.25*frs7)*hrs7;
Ezsq:=0.25*Z2+0.16667*Z4+0.125*Z6+0.10*Z8+0.083333*Z10+1/14*Z12+0.0625*Z14;
RG:=sqrt(Eysq+Ezsq);      a0:=vp[1,n+1];          a2:=a0*a0;
Z0:=Eysq-Ezsq;            Ast:=sign(Z0)*sqrt(abs(Z0));
X1:=bs1;     X3:=bs3+0.75*ds3;        X5:=bs5+0.75*ds5+0.625*fs5;
X7:=bs7+0.75*ds7+0.625*fs7+0.546875*hs7;
V1:=brs1;    V3:=brs3+0.25*drs3;      V5:=brs5+0.25*drs5+0.125*frs5;
V7:=brs7+0.25*drs7+0.125*frs7+0.078125*hrs7;
I4:=0.5*asq*(b1/2+asq*((b3+0.75*d3)/3+asq*((b5+0.75*d5)/4+asq*b7/5)));
I5:=0.5*asq*(br1/2+asq*((br3+0.25*dr3)/3+asq*((br5+0.25*dr5)/4+asq*br7/5)));
intxd:=(oddbits[48]-oddbits[47])/shifno;        weight:=oddbits[10+jj];
            for shif:=1 step 1 until shifno do
                begin xd:=oddbits[47]+shif*intxd;
                      P1[app,ang,shif]:=sqrt(Eysq+Ezsq+2*xd*a0*(I4+I5)+0.5*xd*xd a2*a*a);
                      Y2:=Eysq-Ezsq+2*xd*a0*(I4-I5);
                      P2[app,ang,shif]:=sign(Y2)*sqrt(abs(Y2));
                end shif;
            xopt[1,app,ang]:=best focus(a,V);
            xopt[2,app,ang]:=if abs((I4-I5)/Eysq)<⏨-10 then 0.0 else (Eysq-Ezsq)/(2*a0*(I5-I4));
```

```
                    if jj<angno then goto more;
finish:             end apno loop;
        for app:=1 step 1 until apno do
            begin if app=1 then top of form else print $$15??;     print $APERTURE?,oddbits[40+app],$$1??;
                print $    XDASH                                                    VARIOUS ANGLES$1??;
                for shif:=1 step 1 until shifno do
                    begin print prefix($$s??), scaled(3),$$1??,oddbits[47]+shif*intxd,$$s4 ?;
                        for ang:=1 step 1 until angno do print scaled(3),prefix($ ?),P1[app,ang,shif],
                                P2[app,ang,shif],$$s??;
                    end shif;
                print $$11?BEST FOCUS PLANES$1?SPREAD?;
                for ang:=1 step 1 until angno do print xopt[1,app,ang];
                print $$1?ASTIGMATISM?;            for ang:=1 step 1 until angno do print xopt[2,app,ang];
            end print out;
        clear(OPTXD);              code 30<9BIP> :  20 i;  OPTXD[1]:=i;
        OPTXD[2]:=3;               OPTXD[8]:=volume;       open(OPTXD);
        OPTXD[10]:=2*apno*angno;            OPTXD[13]:=address(xopt);        out40(OPTXD);
        end penultimate;           print $$1??;
        if ng(3) then begin finalise;   code 30<9P8A> : 20 i;   chain(i);   end ng(8) else
        if ng(1) then begin code 30 <9P1A> : 20 i;      finalise;   chain(i) end;
end program P5B;
```

```
VST,30-11-74;
begin      integer i,j,n,apno,angno,freqno,shifno,volume;   real lambda,e,vbk;
           integer array ODDS,OPTXD[1:15];    comment CBS:; array oddbits[1:200];

           procedure chain(program);  value program;  integer program;
           comment calls & enters a program dumped on disk by AFDUMP. Program
                  is the program name in rappacked form;
                code 30 9 : 46 1,
                     40 0, : 26 7886
                     72 256 : 30 <+21>
                     72 6659 : 30 <-47>
                     20 11, : 75 6656
                     32 11, : 45 4,
                     26 7913 : 30 program
                     50 24 : 72 6657
                     54 12 : 72 6657
                     54 12 : 72 6657
                     44 8166 : 00 0
                     00 0 : 00 0;
           comment end of chain;

procedure adjust(forward,h,rho,e,c);    value forward,h,rho,e;   real h,rho,e;
           integer forward;   array c;
      begin comment This procedure modifies the Buchdahl retardation coefficients to allow expression of the
           retardation function in terms of coordinates (y,z) with origin shifted by 0.375/(Fno*Fno). This
           shift corresponds to that needed (according to a third order approximation to ymax) to approximate
           the pupil on the reference sphere by a circle (or ellipse). Also the modifications required to
           reduce the pupil to a circle are incorporated into the coefficient expressions. Details
           of the approximations are given in chapter three of my thesis. The parameter forward is 0 for the
           adjustment to take place, and is 1 if the original values of the coefficients are to be restored.
            rho is the radius of the paraxial exit pupil, and c is the array containing the original coefficients.;

           array store[1:28];    integer i,j;          real a,a2,a3,a4,a5,a6,a7;    switch L:=jump;
           a:=1.5*rho*rho/(e*e);  a2:=a*a;  a3:=a2*a;  a4:=a3*a;  a5:=a4*a;  a6:=a5*a;  a7:=a6*a;
           if forward=0 then
                  begin for i:=1 step 1 until 28 do store[i]:=c[i] end
           else
                  begin for i:=1 step 1 until 28 do c[i]:=store[i];     goto jump end;
```

```
      c[28]:=c[28]+2*a*(c[27]+c[24])+3*a2*(c[26]+c[23])+4*a3*(c[25]+c[22]+c[20])
              +5*a4*(c[21]+c[19])+6*a5*(c[18]+c[17])+7*a6*c[16]+8*a7*c[15]
              -0.9*c[14];
      c[27]:=c[27]+2*a*c[24]+a2*(6*c[25]+5*c[22]+4*c[20])+a3*(9*c[21]+8*c[19])
              +a4*(13*c[18]+12*c[17])+18*a5*c[16]+24*a6*c[15]
              -1.8*(c[13]+c[11])+0.81*(c[3]+c[4]);
      c[26]:=c[26]+4*a*c[25]+2*a*c[22]+a2*(7*c[21]+4*c[19])+4*a3*(3*c[18]+2*c[17])
              +30*a4*c[16]+32*a5*c[15]-2.7*c[12]+0.81*c[2];
      c[25]:=c[25]+2*a*c[21]+4*a2*c[18]+8*a3*c[16]+16*a4*c[15]+3.24*c[1];
      c[24]:=c[24]+a*c[23]+a2*(c[22]+2*c[20])+a3*(c[21]+2*c[19])+a4*(2*c[18]+3*c[17])
              +3*a5*c[16]+4*a6*c[15];
      c[23]:=c[23]+2*a*(c[22]+2*c[20])+3*a2*(c[21]+2*c[19])+4*a3*(2*c[18]+3*c[17])
              +15*a4*c[16]+24*a5*c[15]-0.9*c[10];
      c[22]:=c[22]+a*(3*c[21]+4*c[19])+a2*(10*c[18]+12*c[17])+24*a3*c[16]+48*a4*c[15]
              -1.8*(2*c[8]+c[9])+1.62*c[1];
      c[21]:=c[21]+4*a*c[18]+12*a2*c[16]+32*a3*c[15]-3.6*c[7];
      c[20]:=c[20]+a*c[19]+a2*(c[18]+3*c[17])+3*a3*c[16]+6*a4*c[15];
      c[19]:=c[19]+2*a*(c[18]+3*c[17])+9*a2*c[16]+24*a3*c[15]-0.9*c[7];
      c[18]:=c[18]+6*a*c[16]+24*a2*c[15]-5.4*c[6];
      c[17]:=c[17]+a*c[16]+4*a2*c[15];
      c[16]:=c[16]+8*a*c[15];

      c[14]:=c[14]+2*a*(c[11]+c[13])+3*a2*(c[10]+c[12])+4*a3*(c[8]+c[9])+5*a4*c[7]
              +6*a5*c[6]-0.9*c[5];
      c[13]:=c[13]+a*(3*c[12]+2*c[10])+a2*(5*c[9]+4*c[8])+8*a3*c[7]+12*a4*c[6]-1.8*(c[3]+c[4]);
      c[12]:=c[12]+2*a*c[9]+4*a2*c[7]+8*a3*c[6]-1.8*c[2];
      c[11]:=c[11]+a*c[10]+a2*(c[9]+2*c[8])+2*a3*c[7]+3*a4*c[6];
      c[10]:=c[10]+2*a*(c[9]+2*c[8])+6*a2*c[7]+12*a3*c[6]-0.9*c[2];
      c[9]:=c[9]+4*a*c[7]+12*a2*c[6]-3.6*c[1];
      c[8]:=c[8]+a*c[7]+3*a2*c[6];
      c[7]:=c[7]+6*a*c[6];

      c[5]:=c[5]+2*a*(c[3]+c[4])+3*a2*c[2]+4*a3*c[1];
      c[4]:=c[4]+2*a*c[2]+4*a2*c[1];
      c[3]:=c[3]+a*c[2]+2*a2*c[1];
      c[2]:=c[2]+4*a*c[1];
jump:
   end adjust;
```

```
         volume:=(if m(18) then 19 else 8);
         clear(ODDS);        code 30<9ODDS>:20i;      ODDS[1]:=i;
         ODDS[2]:=3;         ODDS[7]:=1;              ODDS[8]:=volume;
         open(ODDS);         ODDS[10]:=size(oddbits);          ODDS[13]:=address(oddbits);
         in40(ODDS);         n:=entier(oddbits[1]);   lambda:=oddbits[22];
         e:=oddbits[34];     vbk:=oddbits[35];        angno:=entier(oddbits[5]+0.01);
         freqno:=entier(oddbits[16]+0.01);   apno:=entier(oddbits[40]+0.01);
         shifno:=entier(oddbits[46]+0.01);

begin    integer m,p,q,ap,ang,freq,shif;
         integer array ABCO,TAB[1:15];
         real AP,AP2,AP4,AP6,AP8,VV,hk,h2,h4,h6,spf,s2,s4,s6,s8,xd,E00,E20,E22,E40,E42,E44,E60,
         E62,E64,E66,E80,E82,E84,E86,E88,E100,E102,E104,E106,E120,E122,E124,E140,E142,z1,z2,
         pi1,pi2,pi3,pi4,pi5,sig1,sig2,sig3,sig4,sig5,sig6,sig7,sig8,sig9,tau1,tau2,tau3,tau4,
         tau5,tau6,tau7,tau8,tau9,tau10,tau11,tau12,tau13,tau14,a11,a31,a51,a71,a32,a52,a72,a73,
         b11,b31,b51,b71,b52,b72,A11,A31,A51,A71,A33,A53,B22,B42,B44,B62,C00,C20,C40,C22,C42,
         C62,C44,D11,D31,D51,D71,D33,D53,x,y,z,Fno,C60,xd1,xd2,xd3,P3,const,P1sqbr,P2sqbr,a20,
         a22,a40,a42,a44,a60,a62,a80,bb11,bb31,bb33,bb51,bb53,bb71;
comment CBS:;  array table[1:7,0:4,1:50],apert[1:apno],angles[1:angno],freqs[1:freqno],coeff[1:30];
         comment CBS:;    array xshift[1:2,1:shifno],uppx,lowx[1:2,1:apno,1:angno],
                   optxd[1:2,1:5,1:5,1:5],xoptvar[1:5,1:5],V[1:2,1:apno,1:angno,1:freqno,1:shifno],
                   VAR[1:apno,1:angno,1:shifno];
         switch L:=L1,L2,L3,out,finish;

         real procedure Eij(i,j,s,a);   value i,j,s,a;    real s,a;    integer i,j;
               begin z:=50*s/a;   p:=entier(z+₁₀-5);
                   if p<0 then print lineprinter, $$1? * TOO LOW A SPATIAL FREQUENCY ?, scaled(4), z,
                   $    SHOULD BE GREATER THAN ?,0.04*apert[ap]/(e*lambda),$$1??;
                   if p<1 then begin z:=1.0;  p:=1;  end;   x:=table[idiv2,jdiv2,p];
                   y:=table[idiv2,jdiv2,p+1];   Eij:=(x+(y-x)*(z-p))*a↑(i+2);
               end Eij interpolation;

         prefix($$s4??);     scaled(5);  digits(2);  lineprinter;          reader(2);
         clear(ABCO);        clear(TAB);
         comment Program P1A will have computed the ObarT coefficients and used those in P5A to
         determine the spot diagram quantities.P5A will also have determined the pupil periphery.
            Program P8A will use those coefficents to obtain the W-coefficients by transformation
         and then the relations of O.A.C. VII (6.5), (6.6) and (6.7) to obtain the deformation
         coefficients. Using (3.4) these are then converted to retardation coefficients.
                       The following program uses the retardation coefficients and F-number provided
         by P8A in the ABCO file to determine the wavefront difference function variance.   This
```

is evaluated for a prescribed set of apertures for a set of field angles and spatial frequencies and image planes specified through reader two input. The optimum image plane using minimization of this variance is also calculated for prescribed aperture, field angle and frequency.;

```
code 30<9ABS>:20i;      ABCO[1]:=i;             ABCO[2]:=3;
ABCO[7]:=1;             ABCO[8]:=volume;        open(ABCO);
ABCO[10]:=size(coeff);  ABCO[13]:=address(coeff);               in40(ABCO);
code 30<9EIJ>:20i;      TAB[1]:=i;  TAB[2]:=3;
TAB[7]:=1;  TAB[8]:=volume;         open(TAB);
TAB[10]:=size(table);   TAB[13]:=address(table);                in40(TAB);
for i:=1 step 1 until apno do aport[i]:=oddbits[40+i];    vbk:=oddbits[35];
Fno:=oddbits[36];          e:=oddbits[34];        const:=6.283152*3.1415926/(lambda*lambda);
clear(V);   clear(uppx);               clear(lowx);             clear(optxd);
for i:=1,2 do for ap:=1 step 1 until apno do for ang:=1 step 1 until angno do
      begin uppx[i,ap,ang]:=oddbits[48];         lowx[i,ap,ang]:=oddbits[47]  end;
for freq:=1 step 1 until freqno do
      begin if ng(4) then read freqs[freq] else freqs[freq]:=oddbits[16+freq] end;
for ang:=1 step 1 until angno do angles[ang]:=oddbits[5+ang];
if ng(2) then for ap:=1 step 1 until apno do for ang:=1 step 1 until angno do
for i:=1,2 do read lowx[i,ap,ang],uppx[i,ap,ang];
for ap:=1 step 1 until apno do
      begin AP:=aport[ap];
            for freq:=1 step 1 until freqno do
                begin spf:=freqs[freq]*0.5*e*lambda;
                      if spf>AP then goto L1;              s2:=spf*spf;
                      s4:=s2*s2;  s6:=s4*s2;  s8:=s6*s2;
                      for ang:=1 step 1 until angno do
                          begin VV:=angles[ang];           hk:=e*vbk*tan(VV*0.01745329);
                                AP:=AP*(1.0-0.5*(AP*AP+hk*hk)/(e*e));
                                if spf>AP then goto L1;             AP2:=AP*AP;
                                AP4:=AP2*AP2;                       AP6:=AP4*AP2;
                                a11:=arccos(spf/AP);
                                E00:=0.5*AP*(a11*AP-spf*sin(a11));
                                E20:=Eij(2,0,spf,AP);    E22:=Eij(2,2,spf,AP);
                                E40:=Eij(4,0,spf,AP);    E42:=Eij(4,2,spf,AP);
                                E44:=Eij(4,4,spf,AP);    E60:=Eij(6,0,spf,AP);
                                E62:=Eij(6,2,spf,AP);    E64:=Eij(6,4,spf,AP);
                                E66:=Eij(6,6,spf,AP);    E80:=Eij(8,0,spf,AP);
                                E82:=Eij(8,2,spf,AP);    E84:=Eij(8,4,spf,AP);
```

```
E86:=Eij(8,6,spf,AP);    E88:=Eij(8,8,spf,AP);
E100:=Eij(10,0,spf,AP); E102:=Eij(10,2,spf,AP);
E104:=Eij(10,4,spf,AP); E106:=Eij(10,6,spf,AP);
E120:=Eij(12,0,spf,AP); E122:=Eij(12,2,spf,AP);
E124:=Eij(12,4,spf,AP); E140:=Eij(14,0,spf,AP);
E142:=Eij(14,2,spf,AP);
h2:=hk*hk;  h4:=h2*h2;  h6:=h4*h2;
if not ng(17) then adjust(0,hk,apert[ap],e,coeff);
pi1:=coeff[1];          pi2:=coeff[2];          pi3:=coeff[3];
pi4:=coeff[4];          pi5:=coeff[5];
sig1:=coeff[6];         sig2:=coeff[7];         sig3:=coeff[8];
sig4:=coeff[9];         sig5:=coeff[10];        sig6:=coeff[11];
sig7:=coeff[12];        sig8:=coeff[13];        sig9:=coeff[14];
tau1:=coeff[15];        tau2:=coeff[16];        tau3:=coeff[17];
tau4:=coeff[18];        tau5:=coeff[19];        tau6:=coeff[20];
tau7:=coeff[21];        tau8:=coeff[22];        tau9:=coeff[23];
tau10:=coeff[24];       tau11:=coeff[25];       tau12:=coeff[26];
tau13:=coeff[27];       tau14:=coeff[28];
a11:=pi3*h2+sig6*h4+tau10*h6;
a31:=pi1+sig3*h2+tau6*h4;               a32:=pi4+sig8*h2+tau13*h4;
a51:=sig1+tau3*h2;                      a52:=sig4+tau8*h2;
a71:=tau1;  a72:=tau4;  a73:=tau11;
b11:=pi5*h2+sig9*h4+tau14*h6;
b31:=pi2+sig5*h2+tau9*h4;               b51:=sig2+tau5*h2;
b52:=sig7+tau12*h2;      b71:=tau2;  b72:=tau7;
A11:=a11+2*a31*s2+3*a51*s4+4*a71*s6;
A31:=2*a31+6*a51*s2+12*a71*s4+a52*h2+2*a72*h2*s2;
A51:=2*a72*h2+3*a51+12*a71*s2;         A71:=4*a71;
A33:=16*a71*s4-(a52*h2+2*a72*h2*s2-4*a51*s2);
A53:=-2*(a72*h2-8*a71*s2);
B22:=(b31+2*b51*s2+3*b71*s4)*hk;
B42:=(2*b51+b72*h2+6*b71*s2)*hk;
B62:=3*hk*b71;           B44:=(4*b71*s2-b72*h2)*hk;
C00:=(b11+b31*s2+b51*s4+b71*s6+b52*h2*s2+b72*h2*s4)*hk;
C20:=(b31+2*b51*s2+3*b71*s4)*hk;
C40:=(b51+3*b71*s2)*hk;                C60:=b71*hk;
C22:=(2*b31+8*b51*s2+18*b71*s4+3*b52*h2+9*b72*h2*s2)*hk;
C42:=(4*b51+24*b71*s2+3*b72*h2)*hk;
C62:=6*b71*hk;           C44:=(8*b71*s2+2*b72*h2)*hk;
D11:=(a11+2*a31*s2+3*a51*s4+4*a71*s6+a32*h2+2*a52*h2*s2+3*a72*h2*s4
           +2*a73*h4*s2)*2;
D31:=(2*a31+6*a51*s2+12*a71*s4+a52*h2+4*a72*h2*s2)*2;
```

```
D51:=(3*a51+12*a71*s2+a72*h2)*2;
D71:=8*a71;
D33:=2*(4*a51*s2+16*a71*s4+a52*h2+6*a72*h2*s2+2*a73*h4);
D53:=2*(16*a71*s2+a72*h2);

if freq=1 then
    begin a20:=a11;    a22:=a32*h2;   a40:=a31;      a42:=h2*a52;
          a44:=a73*h4; a60:=a51;      a62:=a72*h2;   a80:=a71;
          bb11:=b11;   bb31:=b31;     bb33:=b52*h2;
          bb51:=b51;   bb53:=b72*h2;  bb71:=b71;
    end freq=1;

for shif:=1 step 1 until shifno do
    begin comment segment:;          real dum1;
          xd1:=lowx[1,ap,ang]+(uppx[1,ap,ang]-lowx[1,ap,ang])*shif/shifno;
          xd2:=lowx[2,ap,ang]+(uppx[2,ap,ang]-lowx[2,ap,ang])*shif/shifno;
          xd3:=oddbits[47]+(oddbits[48]-oddbits[47])*shif/shifno;
          A11:=A11+0.5*xd1/(e*e);                D11:=D11+xd2/(e*e);
          V[1,ap,ang,freq,shif]:=16*s2/E00*(A11*A11*E22
          +2*A11*(A31*E42+A71*E82)+(A31*A31+2*A11*A51)*E62
                +(A51*A51+2*A31*A71)*E102+2*(A31*A51*E82+A51*A71*E122)+A71*A71*E142
                +2*A33*(A11*E44+A31*E64+A51*E84+A71*E104)+2*A53*(A11*E64
                +A31*E84+A51*E104+A71*E124)+2*A53*A33*E86+A53*A53*E106
                +A33*A33*E66+B22*B22*(E42-E44)+2*B22*B42*(E62-E64)+B42*B42*(E82-E84)
                +2*B42*B62*(E102-E104)+B62*B62*(E122-E124)+2*B22*B44*(E64-E66)+
                2*B42*B44*(E84-E86)+2*B62*B44*(E104-E106)+B44*B44*(E86-E88));
          P3:=C00*E00+C20*E20+C40*E40+C60*E60+C22*E22+C42*E42+C62*E62+C44*E44;
          V[2,ap,ang,freq,shif]:=4*s2/E00*(-P3*P3/E00+C00*C00*E00+C20*C20*E40+C40*C40*E80
                +C60*C60*E120+C22*C22*E44+C42*C42*E84+C62*C62*E124+C44*C44*E88
                +2*C00*(C20*E20+C40*E40+C60*E60+C22*E22+C42*E42+C62*E62+C44*E44)
                +2*C20*(C40*E60+C60*E80+C22*E42+C42*E62+C62*E82+C44*E64)+2*C40*
                (C60*E100+C22*E62+C42*E82+C62*E102+C44*E84)+2*C60*(C22*E82+C42*E102
                +C62*E122+C44*E104)+2*C22*(C42*E64+C62*E84+C44*E66)+2*C42*(C62*E104
                +C44*E86)+2*C62*C44*E106+D11*D11*E22+D31*D31*E62+D51*D51*E102+D71*D71*E142
                +D33*D33*E66+D53*D53*E106+2*D11*(D31*E42+D51*E62+D71*E82+D33*E44
                +D53*E64)+2*D31*(D51*E82+D71*E102+D33*E64+D53*E84)+2*D51*(D71*E122
                +D33*E84+D53*E104)+2*D71*(D33*E104+D53*E124)+2*D33*D53*E86);
          if freq=1 then
                begin a20:=a20+0.5*xd3/(e*e);
```

```
                                        P2sqbr:=h2*AP2*(bb11*bb11/4+bb11*bb31/3*AP2+(bb31*bb31+2*bb11*bb51)*AP4/8
                                          +(bb11*bb71+bb31*bb51)*AP6/5+(bb51*bb51+2*bb31*bb71)*AP4*AP4/12
                                          +bb51*bb71/7*AP6*AP4+bb71*bb71/16*AP6*AP6+bb11*bb33/4*AP2
                                          +3*(bb11*bb53+bb31*bb33)/16*AP4+3*(bb31*bb53+bb51*bb33)/20*AP6
                                          +(bb51*bb53+bb71*bb33)/8*AP4*AP4+3*bb71*bb53/28*AP6*AP4
                                          +5*bb33*bb33/64*AP4+bb33*bb53/8*AP6+5*bb53*bb53*AP4*AP4/96);
                                        P1sqbr:=AP4*(a20*a20/12+a20*a40*AP2/6+(4*a40*a40/45+3*a20*a60/20)*AP4
                                          +(2*a20*a80/15+a40*a60/6)*AP6+(9*a60*a60/112+16*a40*a80/105)*AP4*AP4
                                          +3*a60*a80/20*AP6*AP4+16*a80*a80/225*AP6*AP6+a20*a22/12
                                          +(a20*a42+a22*a40)/12*AP2+(3*a20*a62/40+4*a40*a42/45+3*a60*a22/40)*AP4
                                          +(a40*a62/12+a60*a42/12+a80*a22/15)*AP6+(9*a60*a62/112+8*a80*a42/105)*
                                          AP4*AP4+3*a80*a62/40*AP6*AP4+a22*a22/16+(a20*a44+5*a22*a42/3)/16*AP2
                                          +(17*a42*a42/360+7*a22*a62/80+a40*a44/15)*AP4+(a42*a62/12
                                          +a60*a44/16)*AP6+(17*a62*a62/448+2*a80*a44/35)*AP4*AP4+3*
                                          a22*a44/32*AP2+a42*a44/12*AP4+7*a62*a44/96*AP6+3/128*a44*a44*AP4*AP4);
                                        VAR[ap,ang,shif]:=P1sqbr+P2sqbr;
                                        a20:=a20-0,5*xd3/(e*e);
                                    end freq=1;        D11:=D11-xd2/(e*e);
                                A11:=A11-0,5*xd1/(e*e);
                            end shif;
                                if freq=1 then xoptvar[ap,ang]:=-e*e*(2*a20+a22+(2*a40+a42+0,75*a44)*AP2
                                                                   +(1.8*a60+0.9*a62)*AP4+1.6*a80*AP6);
                        optxd[1,ap,ang,freq]:=-2*e*e/E22*(A11*E22+A31*E42+A51*E62+A71*E82+A33*E44+A53*E64);
                        optxd[2,ap,ang,freq]:=-e*e/E22*(D11*E22+D31*E42+D51*E62+D71*E82+D33*E44+D53*E64);
                        if not ng(17) then adjust(1,hk,apert[ap],e,cooff);         AP:=apert[ap];
                    end angle loop;
1:                  end frequency loop;
                end aperture loop;
                for ap:=1 step 1 untilapno do
                    begin top of form;           print $APERTURE ?,apert[ap],$$1??;
                        print $   XDASH                        SAGITTAL                    TANGENTIAL$1??;
                        for ang:=1 step 1 until angno do
                            begin print $$1?ANGLE ?,angles[ang],$$1??;
                                for shif:=1 step 1 until shifno do
                                begin for i:=1,2 do
                                    begin ifi=2 then
                                        begin j:=pointer;  for p:=1 step 1 until 63-j do print $$s??;  end;
                                print prefix($ ?),freepoint(5),lowx[i,ap,ang]+(uppx[i,ap,ang]
                                        -lowx[i,ap,ang])*shif/shifno;
                                        for freq:=1 step 1 until freqno do print prefix($$s??),
                                              scaled(3),1.0-V[i,ap,ang,freq,shif]*const;
```

```
                        nd i;           print $$1??;
                    end shif;   print $$1??;

                end ang;
        end freq;
    top of form;           freepoint(5);          prefix($$ss??);
    print $$s25?*  *  *  FUNCTIONS S1 and S2  *  *  * $11??;
    for ap:=1 step 1 until apno do
        begin if ap>1 then top of form;    print $APERTURE   ?,apert[ap],$   FNO ?,0.5*e/apert[ap];
            for ang:=1 step 1 until angno do
                begin print $$13?ANGLE  ?,angles[ang];
                    print $$1s24?XDASH$s4?FREQ$s7?S1$s7?S2?;
                    AP:=apert[ap];          hk:=tan(angles[ang]*0.01745329);
                    AP:=AP*(1-(0.75*AP*AP-0.5*hk*hk)/(e*e));
                    for shif:=1 step 1 until shifno do
                        begin if shif>0 then print $$1??;
                            print $$1s20??,oddbits[47]+(oddbits[48]-oddbits[47])*shif/shifno;
                            for freq:=1 step 1 until freqno do
                                begin if freq>1 then print $$1s29??;
                                    a11:=1-V[1,ap,ang,freq,shif]*const;
                                    a20:=1-V[2,ap,ang,freq,shif]*const;
                                    a31:=a11*a20;          spf:=0.5*e*freqs[freq]*lambda;
                                    print freqs[freq],sign(a31)*sqrt(abs(a31)),a11-a20;
                                end freq;
                        end shif;
                end ang;
        end ap;
            top of form;
            for ap:=1 step 1 until apno do
                begin if ap=1 then
                    begin top of form; print $OPTIMUM IMAGE PLANES, VARIOUS ANGLES ACROSS,
                              APERTURE SETS DOWNWARDS$1?? end
                        else print $$111??;
                    print $FREQUENCY   SAGITTAL BEST PLANES $1?              TANGENTIAL BEST PLANES?;
                    for freq:=1 step 1 until freqno do
                        begin print scaled(3),$$1??,prefix($?),freqs[freq];
                            for i:=1,2 do
                                begin print $$ss??;  for ang:=1 step 1 until angno do
                                    print prefix(S ?),scaled(3),optxd[i,ap,ang,freq] end i;
                        end freq;
                end ap;
    print $$15?FRACTIONAL SPATIAL FREQUENCIES USED$1??;
    for ap:=1 step 1 until apno do
```

```
          print $$15?FRACTIONAL SPATIAL FREQUENCIES USED$1??;
          for ap:=1 step 1 until apno do
              begin print $$1?APERTURE?,apert[ap];
                  for freq:=1 step 1 until freqno do print 0.5*o*freqs[freq]*lambda/apert[ap];
              end ap loop;
                  print top of form,$VARIANCE OF WAVEFRONT RETARDATION$11?    APERTURE  XDASH     ANGLES$1??;
                  for ap:=1 step 1 until apno do
                      begin scaled(4);  prefix($$ss??);
                          print $$1??,apert[ap],$$ss??;
                          for shif:=1 step 1 until shifno do
                              begin if shif>1 then print $$1s14??;
                                  print oddbits[47]+(oddbits[48]-oddbits[47])*shif/shifno,$$s??;
                                  for ang:=1 step 1 until angno do
                                      print VAR[ap,ang,shif]/(lambda*lambda);
                              end shif;
                          print $$11?BEST IMAGE PLANES,  VARIOUS ANGLES$1??;
                          for ang:=1 step 1 until angno do print xoptvar[ap,ang];     print $$111??;
                      end ap;
                  clear(OPTXD);              code 30<9BIP> :  20 i;             OPTXD[1]:=i;
                  OPTXD[2]:=3;               OPTXD[7]:=1;           OPTXD[8]:=volume;        open(OPTXD);
                  OPTXD[10]:=size(optxd);          OPTXD[13]:=address(optxd);           out40(OPTXD);
                  OPTXD[10]:=size(xoptvar);        OPTXD[13]:=address(xoptvar);
                  out40(OPTXD);
      end main block;
      erase(ODDS);       clear(ODDS);           code 30<9ODDS> :  20 i;             ODDS[1]:=i;
      ODDS[2]:=3;        ODDS[8]:=volume;       open(ODDS);
      ODDS[10]:=size(oddbits);     ODDS[13]:=address(oddbits);        out40(ODDS);
      if ng(12) then
          begin finalise;    code 30<9P12> :  20 i;    chain(i)  end ng(12);
          if ng(13) and not ng(12) then
              begin finalise;    code 30<9P13> :  20 i;    chain(i)   end ng(13);
      if ng(1) then
          begin code 30 <9P1A> : 20 i;    finalise;  chain(i) end ng(1);
  end of program;
```

P12, 29-11-74;

```
begin integer i,j,m,n,angno,apno,freqno,shifno,volume,NMESH;            integer array ODDS,ABCO[1:15];
comment CBS:;array oddbits[1:200];  array cooff[1:30];

          procedure chain(program);  value program;  integer program;
          comment calls & enters a program dumped on disk by AFDUMP. Program
               is the program name in rappacked form;
             code 30 9 : 46 1,
                40 0, : 26 7886
                72 256 : 30 <+21>
                72 6659 : 30 <-47>
                20 11, : 75 6656
                32 11, : 45 4,
                26 7913 : 30 program
                50 24 : 72 6657
                54 12 : 72 6657
                54 12 : 72 6657
                44 8166 : 00 0
                00 0 : 00 0;
          comment end of  chain;

          clear(ODDS);              clear(ABCO);              clear(oddbits);
          volume:=if ng(18) then 19 else 8;
          code 30<9ODDS> :  20 i;          ODDS[1]:=i;
          code 30<9ABS> : 20 i;            ABCO[1]:=i;
          ABCO[2]:=ODDS[2]:=3;    ABCO[7]:=ODDS[7]:=1;
          ABCO[8]:=ODDS[8]:=volume;        open(ABCO);          open(ODDS);
          ABCO[10]:=size(cooff);           ABCO[13]:=address(cooff);        in40(ABCO);
          ODDS[10]:=size(oddbits);         ODDS[13]:=address(oddbits);      in40(ODDS);

          angno:=entier(oddbits[5]+0.1);     apno:=entier(oddbits[40]+0.1);
          freqno:=entier(oddbits[16]+0.1);   NMESH:=entier(oddbits[23]+0.1);
          shifno:=entier(oddbits[46]+0.1);
          if ng(3) then read reader(2),apno,angno,shifno;
          if angno*apno*freqno*(shifno+1)>2000 then shifno:=2000 div (apno*angno*freqno)-1;
     begin real w20,w11,w40,w31,w22,w60,w51,w42,w33,w80,w71,w62,w53,w44,uppx,lowx,intx,xd,low,upp,delz,
          dely,a,b,real,imag,mod,phi,sumr,sumi,x,y,z,A,B,C,D,spf,s,t,Fno,lambda,e,vbk,hk,h2,h4,h3,h6,
          AP,V,arg,oc,int;
          comment CBS:;array REAL,IMAG,MOD,PHASE[1:2,1:apno,1:angno,0:shifno,1:freqno],xshift[1:2,1:apno,
          1:angno,0:shifno],TH[1:apno,1:freqno],aport[1:apno],angles[1:angno],freqs[1:freqno];
          integer ap,ang,freq,shif;
          switch L:=jump;
```

```
real procedure W(x,y,s,t);
value x,y,s,t;          real x,y,s,t;
    begin real sq,xs,y2,yt;
        xs:=x+s;      yt:=y+t;       y2:=yt*yt;          sq:=xs*xs+y2;
        W:=(((w80*sq+w60)*sq+w40)*sq+w20)*sq + (((w71*sq+w51)*sq+w31)*sq)*yt
              + ((w62*sq+w42)*sq+w22)*y2 + (w53*sq+w33)*y2*yt+ w44*y2*y2
    end W;

real procedure Wy(x,y,s,t);       value x,y,s,t;      real x,y,s,t;
    begin real g1,g2,g3,g4,g5,g6,g7,g8,x2,xs,sum;      integer i;
        xs:=x+s;      x2:=xs*xs;     y:=y+t;
        g1:=((w71*x2+w51)*x2+w31)*x2;
        g2:=((4*w80*x2+w62+w60*3)*x2+w42+2*w40)*x2+w20+w22;
        g3:=(w71*3*x2+w53+2*w51)*x2+w33+w31;
        g4:=(w80*6*x2+2*w62+3*w60)*x2+w44+w42+w40;
        g5:=3*w71*x2+w51+w53;
        g6:=4*x2*w80+w60+w62;
        g7:=w71;                         g8:=w80;
        Wy:=(((((8.0*g8*y+7.0*g7)*y+6.0*g6)*y+5.0*g5)*y+4.0*g4)*y+3.0*g3)*y+
            2.0*g2)*y+g1;
        y:=y-t;
    end procedure;

real procedure Wx(x,y,s,t);       value x,y,s,t;      real x,y,s,t;
    begin real h2,h4,h6,h8,y2,y4,x2;        integer i;
        y:=y+t;       y2:=y*y;        y4:=y2*y2;          x:=x+s;
        h8:=w80;     h6:=4.0*w80*y2+w71*y+w60;
        h4:=6.0*w80*y4+3.0*w71*y2*y+(w62+3.0*w60)*y2+w51*y+w40;
        h2:=4.0*w80*y4*y2+3.0*w71*y4*y+(2.0*w62+3.0*w60)*y4+(w53+2.0*w51)*y2*y
            +(w42+2.0*w40)*y2+w31*y+w20;
        x2:=x*x;
        Wx:=(((8.0*h8*y2+6.0*h6)*x2+4.0*h4)*x2+2.0*h2)*x;       x:=x-s;     y:=y-t;
    end Wx;

procedure adjust(forward,h,rho,e,c);    value forward,h,rho,e;   real h,rho,e;
        integer forward;    array c;
    begin comment This procedure modifies the Buchdahl retardation coefficients to allow expression of the
        retardation function in terms of coordinates (y,z) with origin shifted by 0.375/(Fno*Fno). This
        shift corresponds to that needed (according to a third order approximation to ymax) to approximate
        the pupil on the reference sphere by a circle (or ellipse). Also the modifications required to
```

reduce the pupil to a circle are incorporated into the coefficient expressions. Details of the approximations are given in chapter three of my thesis. The parameter forward is 0 for the adjustment to take place, and is 1 if the original values of the coefficients are to be restored. rho is the radius of the paraxial exit pupil, and c is the array containing the original coefficients.;

```
array store[1:28];        integer i,j;          real a,a2,a3,a4,a5,a6,a7;     switch L:=jump;
a:=1.5*rho*rho/(c*e);  a2:=a*a;  a3:=a2*a;  a4:=a3*a;  a5:=a4*a;  a6:=a5*a;  a7:=a6*a;
if forward=0 then
      begin for i:=1 step 1 until 28 do store[i]:=c[i] end
else
      begin for i:=1 step 1 until 28 do c[i]:=store[i];    goto jump end;
c[28 c[28]:=c[28]+2*a*(c[27]+c[24])+3*a2*(c[26]+c[23])+4*a3*(c[25]+c[22]+c[20])
              +5*a4*(c[21]+c[19])+6*a5*(c[18]+c[17])+7*a6*c[16]+8*a7*c[15]
              -0.9*c[14];
c[27]:=c[27]+2*a*c[24]+a2*(6*c[25]+5*c[22]+4*c[20])+a3*(9*c[21]+8*c[19])
              +a4*(13*c[18]+12*c[17])+18*a5*c[16]+24*a6*c[15]
              -1.8*(c[13]+c[11])+0.81*(c[3]+c[4]);
c[26]:=c[26]+4*a*c[25]+2*a*c[22]+a2*(7*c[21]+4*c[19])+4*a3*(3*c[18]+2*c[17])
              +20*a4*c[16]+32*a5*c[15]-2.7*c[12]+0.81*c[2];
c[2 c[25]:=c[25]+2*a*c[21]+4*a2*c[18]+8*a3*c[16]+16*a4*c[15]+3.24*c[1];
c[24]:=c[24]+a*c[23]+a2*(c[22]+2*c[20])+a3*(c[21]+2*c[19])+a4*(2*c[18]+3*c[17])
              +3*a5*c[16]+4*a6*c[15];
c[23]:=c[23]+2*a*(c[22]+2*c[20])+3*a2*(c[21]+2*c[19])+4*a3*(2*c[18]+3*c[17])
              +15*a4*c[16]+24*a5*c[15]-0.9*c[10];
c[22]:=c[22]+a*(3*c[21]+4*c[19])+a2*(10*c[18]+12*c[17])+24*a3*c[16]+48*a4*c[15]
              -1.8*(2*c[8]+c[9])+1.62*c[1];
c[21]:=c[21]+4*a*c[18]+12*a2*c[16]+32*a3*c[15]-3.6*c[7];
c[20]:=c[20]+a*c[19]+a2*(c[18]+3*c[17])+3*a3*c[16]+6*a4*c[15];
c[19]:=c[19]+2*a*(c[18]+3*c[17])+9*a2*c[16]+24*a3*c[15]-0.9*c[7];
c[18]:=c[18]+6*a*c[16]+24*a2*c[15]-5.4*c[6];
c[17]:=c[17]+a*c[16]+4*a2*c[15];
c[16]:=c[16]+8*a*c[15];

c[14]:=c[14]+2*a*(c[11]+c[13])+3*a2*(c[10]+c[12])+4*a3*(c[8]+c[9])+5*a4*c[7]
              +6*a5*c[6]-0.9*c[5];
c[13]:=c[13]+a*(3*c[12]+2*c[10])+a2*(5*c[9]+4*c[8])+8*a3*c[7]+12*a4*c[6]-1.8*(c[3]+c[4]);
c[12]:=c[12]+2*a*c[9]+4*a2*c[7]+8*a3*c[6]-1.8*c[2];
c[11]:=c[11]+a*c[10]+a2*(c[9]+2*c[8])+2*a3*c[7]+3*a4*c[6];
c[10]:=c[10]+2*a*(c[9]+2*c[8])+6*a2*c[7]+12*a3*c[6]-0.9*c[2];
c[9]:=c[9]+4*a*c[7]+12*a2*c[6]-3.6*c[1];
```

```
c               c[8]:=c[8]+a*c[7]+3*a2*c[6];
                c[7]:=c[7]+6*a*c[6];

                c[5]:=c[5]+2*a*(c[3]+c[4])+3*a2*c[2]+4*a3*c[1];
                c[4]:=c[4]+2*a*c[2]+4*a2*c[1];
                c[3]:=c[3]+a*c[2]+2*a2*c[1];
                c[2]:=c[2]+4*a*c[1];
jump:
            end adjust;

        real procedure Arg(x,y);      value x,y;      real x,y;
            Arg:= if x>0.0 then arctan(y/x) else if abs(y)>0.0 then sign(y)*1.57079633
                -arctan(x/y) else 3.14159265358;
            comment result in range -pi ..... +pi;

        lineprinter;            scaled(5);  prefix($$s4??);         digits(2);
        ee:=oddbits[34];        vbk:=oddbits[35];       Fno:=oddbits[36];       lambda:=oddbits[22];
        uppx:=oddbits[48];      lowx:=oddbits[47];
        intx:=(if shifno=0 then 0.0 else (uppx-lowx)/shifno);       arg:=6.283152/lambda;
        for m:=1,2 do for ap:=1 step 1 until apno do for ang:=1 step 1 until angno do
        for shif:=0 step 1 until shifno do xshift[m,ap,ang,shif]:=lowx+intx*shif;
        for i:=1 step 1 until apno do apert[i]:=oddbits[40+i];
        for i:=1 step 1 until angno do angles[i]:=oddbits[5+i];
        for i:=1 step 1 until freqno do freqs[i]:=oddbits[16+i];
        if ng(3) then
            begin for i:=1 step 1 until apno do read reader(2),apert[i];
                for i:=1 step 1 until angno do read reader(2),angles[i];
                for m:=1,2 do for ap:=1 step 1 until apno do for ang:=1 step 1 until angno do
                for shif:=0 step 1 until shifno do read reader(2),xshift[m,ap,ang,shif];
            end ng(3);

        clear(REAL);            clear(IMAG);            clear(MOD);             clear(PHASE);
        for ap:=1 step 1 until apno do
        begin AP:=apert[ap];
            for ang:=1 step 1 until angno do
            begin V:=tan(angles[ang]*vbk*0.01745329);           e:=ee/cos(angles[ang]*vbk*0.01745329);
                hk:=ee*vbk*tan(V);       hk:=tan(angles[ang]*0.01745329);
                if ng(20) then print $$1?IMAGE HEIGHT ?,hk;       h2:=hk*hk;
                h3:=h2*hk;   AP:=apert[ap];             h4:=h2*h2;   h6:=h4*h2;
                adjust(0,hk,AP,ee,coeff);
```

```
AP:=AP*(1.0-0.5*(AP*AP+hk*hk)/(e*e));
w20:=coeff[3]*h2+coeff[11]*h4+coeff[24]*h6;
w11:=(coeff[5]*h2+coeff[14]*h4+coeff[28]*h6)*hk;
w40:=coeff[1]+coeff[8]*h2+coeff[20]*h4;
w31:=(coeff[2]+coeff[10]*h2+coeff[23]*h4)*hk;
w22:=coeff[4]*h2+coeff[13]*h4+coeff[27]*h6;
w60:=coeff[6]+coeff[17]*h2;               w51:=(coeff[7]+coeff[19]*h2)*hk;
w42:=coeff[9]*h2+coeff[22]*h4;            w33:=(coeff[12]*h2+coeff[26]*h4)*hk;
w80:=coeff[15];                           w71:=coeff[16]*hk;
w62:=coeff[18]*h2;        w53:=coeff[21]*h3;        w44:=coeff[25]*h4;
for freq:=1 stop 1 until freqno do
begin spf:=freqs[freq]*0.5*e*lambda;
      for shif:=0 stop 1 until shifno do
      begin for m:=1,2 do
               begin real:=imag:=0.0;
                     xd:=xshift[m,ap,ang,shif];          w20:=w20+0.5*xd/(e*e);
                     if (AP-spf)<_10-8 then
                        begin print $$1?TOO HIGH SPATIAL FREQUENCY,    FREQ =?,
                              freqs[freq],$      APERTURE =?,AP,$$1??;
                              goto jump
                        end spf>AP;
                     delz:=(if m=1 then AP-spf else sqrt(AP*AP-spf*spf))/NMESH;
                     if delz<_10-9 then
                        begin print $$1?JUMPED OUT,OTF EFFECTIVELY ZERO?;
                              goto jump
                        end;
                     for j:=1 stop 2 until 2*NMESH do
                        begin z:=(j-NMESH)*delz;
                              x:=z-(if j<NMESH then 1 else -1)*spf;
                              a:=if m=1 then sqrt(AP*AP-x*x) else sqrt(AP*AP-z*z)-spf;
                              dely:=a/NMESH;              int:=2.0*dely;
                              low:=(if m=1 then- a+dely else -(NMESH-1)*dely);
                              upp:=(if m=1 then a else NMESH*dely);
                              sumr:=sumi:=0.0;
                              a:=(if m=1 then spf else 0.0);        b:=spf-a;
                              for y:=low stop int until upp*1.001 do
                                 begin A:=(W(z,y,a,b)-W(z,y,-a,-b))*arg;
                                       B:=(Wx(z,y,a,b)-Wx(z,y,-a,-b))*delz*arg;
                                       C:=(Wy(z,y,a,b)-Wy(z,y,-a,-b))*dely*arg;
                                       A:=sin(A);                 D:=cosino;
                                       if abs(B)<_10-8 then B:=_10-8;
```

```
                                                  if abs(C)<₁₀-8 then C:=₁₀-8;
                                                  x:=sin(B)*sin(C)/(B*C);
                                                  sumr:=sumr+D*x;        sumi:=sumi+A*x;
                                             end y;
                                        real:=real+4.0*dely*delz*sumr;      imag:=imag+4.0*dely*delz*sumi;
                                   end j loop;
                              mod:=sqrt(real*real+imag*imag);     phi:=Arg(real,imag);
                              x:=3.1415926*AP*AP;     mod:=mod/x;          real:=real/x;          imag:=imag/x;
                              REAL[m,ap,ang,shif,freq]:=real;     IMAG[m,ap,ang,shif,freq]:=imag;
                              MOD[m,ap,ang,shif,freq]:=mod;       PHASE[m,ap,ang,shif,freq]:=phi;
                              if abs(V)<₁₀-6 then
                                   begin m:=2;
                                        REAL[m,ap,ang,shif,freq]:=REAL[1,ap,ang,shif,freq];
                                        IMAG[m,ap,ang,shif,freq]:=IMAG[1,ap,ang,shif,freq];
                                        MOD[m,ap,ang,shif,freq]:=MOD[1,ap,ang,shif,freq];
                                        PHASE[m,ap,ang,shif,freq]:=PHASE[1,ap,ang,shif,freq];
               jump:               end V=0;
                         w20:=w20-0.5*xd/(e*e);
                         end m;
                    end shif;
               end freq;
               adjust(1,hk,AP,ee,coeff);
          end ang;
     end ap loop;           e:=ee;

     for ap:=1 step 1 until apno do for freq:=1 step 1 until freqno do
          begin spf:=0.5*freqs[freq]*e*lambda;      AP:=apert[ap];
               if AP<spf then spf:=0.95*AP;         phi:=arccos(spf/AP);
               TH[ap,freq]:=2*AP*(AP*phi-spf*sin(phi))/(3.1415926*AP*AP);
          end ap,freq;
     for m:=1,2 do
     begin if m<3 then top of form;
          print $O.T.F. VALUES?;    if m=1 then print $    SAGITTAL SECTION$11??
               else print $    TANGENTIAL SECTION$11??;
          for ap:=1 step 1 until apno do
          begin if ap>1 then top of form;        print $APERTURE ?,apert[ap],$$1??;
               print $$s6?ANGLE$s10?FREQUENCY$s7?XDASH$s10?MODULUS$s8?PHASE$s11?REAL$s9?IMAGINARY$s6?MOD/IDEAL$1??;
               for ang:=1 step 1 until angno do
               begin print $S1322,angles[ang];
                    for freq:=1 step 1 until freqno do
```

```
AP:=AP*(1.0-0.5*(AP*AP+hk*hk)/(e*e));
w20:=coeff[3]*h2+coeff[11]*h4+coeff[24]*h6;
w11:=(coeff[5]*h2+coeff[14]*h4+coeff[28]*h6)*hk;
w40:=coeff[1]+coeff[8]*h2+coeff[20]*h4;
w31:=(coeff[2]+coeff[10]*h2+coeff[23]*h4)*hk;
w22:=coeff[4]*h2+coeff[13]*h4+coeff[27]*h6;
w60:=coeff[6]+coeff[17]*h2;        w51:=(coeff[7]+coeff[19]*h2)*hk;
w42:=coeff[9]*h2+coeff[22]*h4;     w33:=(coeff[12]*h2+coeff[26]*h4)*hk;
w80:=coeff[15];                    w71:=coeff[16]*hk;
w62:=coeff[18]*h2;      w53:=coeff[21]*h3;      w44:=coeff[25]*h4;
for freq:=1 step 1 until freqno do
begin spf:=freqs[freq]*0.5*e*lambda;
      for shif:=0 step 1 until shifno do
      begin for m:=1,2 do
            begin real:=imag:=0.0;
                  xd:=xshift[m,ap,ang,shif];          w20:=w20+0.5*xd/(e*e);
                  if (AP-spf)<_10-8 then
                     begin print $$1?TOO HIGH SPATIAL FREQUENCY,    FREQ =?,
                           freqs[freq],$      APERTURE =?,AP,$$1??;
                           goto jump
                     end spf>AP;
                  delz:=(if m=1 then AP-spf else sqrt(AP*AP-spf*spf))/NMESH;
                  if delz<_10-9 then
                     begin print $$1?JUMPED OUT,OTF EFFECTIVELY ZERO?;
                           goto jump
                     end;
                  for j:=1 step 2 until 2*NMESH do
                     begin z:=(j-NMESH)*delz;
                           x:=z-(if j<NMESH then 1 else -1)*spf;
                           a:=if m=1 then sqrt(AP*AP-x*x) else sqrt(AP*AP-z*z)-spf;
                           dely:=a/NMESH;          int:=2.0*dely;
                           low:=(if m=1 then- a+dely else -(NMESH-1)*dely);
                           upp:=(if m=1 then a else NMESH*dely);
                           sumr:=sumi:=0.0;
                           a:=(if m=1 then spf else 0.0);       b:=spf-a;
                           for y:=low step int until upp*1.001 do
                              begin A:=(W(z,y,a,b)-W(z,y,-a,-b))*arg;
                                    B:=(Wx(z,y,a,b)-Wx(z,y,-a,-b))*delz*arg;
                                    C:=(Wy(z,y,a,b)-Wy(z,y,-a,-b))*dely*arg;
                                    A:=sin(A);             D:=cosine;
                                    if abs(B)<_10-8 then B:=_10-8;
```

```
    if ng(13) then
        begin finalise;        code 30<9P13> :  20 i;  chain(i);  end ng(13);
    if ng(1) then
        begin finalise;  code 30<9P1A> :  20 i;    chain(i) end;
end of program;
```